

Quiroga, Graciela Alejandra  
Ortiz Banda, Ursula Italia

FACULTAD DE INFORMÁTICA

UNIVERSIDAD NACIONAL DE LA PLATA

*La constancia es la virtud por la que todas las cosas dan su fruto.*

*Arturo Graf*

Dedicado a:

Nuestros padres e hijos.

# ÍNDICE

---

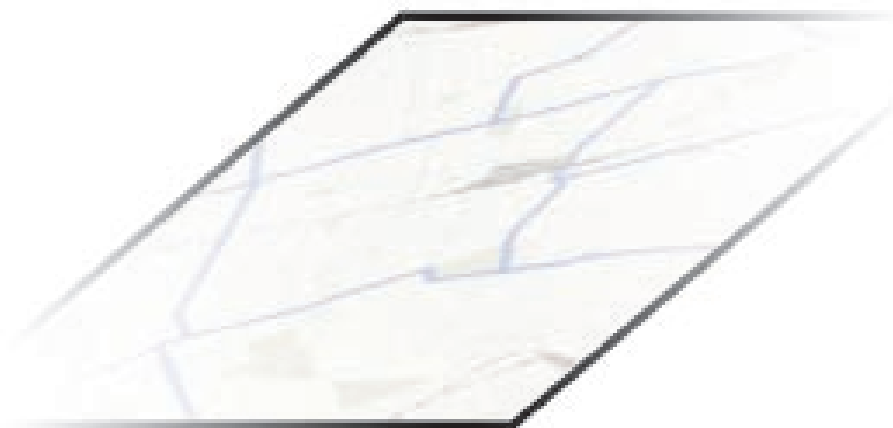
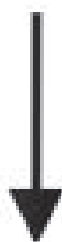
Capítulo I: Introducción .....	4
Prefacio .....	5
Capítulo II: Sistemas de información geográfica .....	8
Introducción .....	9
Mapas .....	9
Base de Datos .....	10
Componentes .....	10
Contexto .....	11
Proceso de Abstracción .....	12
Sistema de Proyección .....	13
Tipos de Datos .....	14
Datos Espaciales .....	14
Representación de los Datos Espaciales .....	15
Formatos de Representación .....	15
Condiciones a Tener en Cuenta .....	21
Datos No Espaciales o Alfanuméricos .....	22
Aplicaciones de SIG .....	23
Beneficios .....	24
Capítulo III: Censos Poblacionales Indicadores Demográficos .....	25
Introducción .....	26
Objetivos de un Censo .....	27
Datos estadísticos necesarios .....	27
Cómo Podemos Utilizar la Información de los Censos.....	27
Definiciones .....	28
La Demografía .....	33
Tamaño de la Población .....	33
Indicadores Demográficos .....	34
Estructura por Edades .....	36
Análisis de Cohorte .....	36
Dinámica Demográfica.....	37
Poblaciones Estables .....	39
Capítulo IV: Definición del Problema .....	45
Introducción .....	46
Definición del Problema .....	46
Características del Problema .....	48
Ejemplificamos el Problema .....	49
Conclusiones Preliminares .....	55
Fundamentación del Paradigma .....	56
Modelo de Datos .....	56

Capítulo V: El modelo .....	60
Introducción .....	61
Análisis de Datos No Espaciales .....	61
Casos de Uso .....	61
Casos de Uso de Alto Nivel .....	61
Caso de Uso expandido .....	63
Fase de Construcción del Modelo Conceptual .....	63
Definición de Asociaciones y Multiplicidad .....	66
Definición de Atributos .....	70
Definición del Comportamiento .....	71
Análisis de Datos Espaciales .....	73
Integración de los datos Espaciales – No Espaciales .....	75
El Modelo: Modelo Conceptual - Modelo Geográfico .....	80
Conclusiones .....	84
Referencia bibliográfica: .....	85



# CAPÍTULO I

## INTRODUCCIÓN





## **PREFACIO**

El punto de partida de este trabajo, fue generar un modelo para la manipulación de censos. Para ello, investigamos sobre Sistemas de Información Geográfica, en adelante SIG, y su aplicación sobre censos, para los cuales fue necesario obtener definiciones precisas, debido a que en el tema muchas palabras como por ejemplo hogar que para nosotros es una palabra común en el contexto de un censo tiene muchos significados.

Además nos fue necesario encontrar un problema real que afecte a la sociedad y sobre el cual la informática tuviese una implicancia significativa, necesitábamos tener en cuenta situaciones cotidianas sobre las que nos vemos obligados a tomar decisiones que tienen una incidencia importante en la población, considerando al individuo como parte activa de una sociedad que actualmente padece muchos conflictos sociales, el tema elegido no es uno solo, sino varios, de los cuales se considero una pequeña parte que podría ser tomada como punto de partida para otras investigaciones.

El problema no es más que un ejemplo de aplicación para mostrar la Arquitectura que se propone, fue elegido por lo que la información representa en el mismo, se caracteriza por necesitar datos precisos acerca de la población para poder proporcionar estadísticas fiables.

Investigamos acerca de las leyes que rigen la divulgación de los datos y de cómo pueden ser publicados, esto fue necesario para entender para qué fines debería ser útil el modelo propuesto, entendimos que el procesamiento de la información es la etapa más importante cuando se realiza un censo. Tener claro como va a ser el mismo y contar con herramientas que permitan combinar la información sin deteriorar su valor práctico es fundamental, ya que las estadísticas que esta va a proporcionar a los distintos usuarios deben ser lo más aproximadas a un porcentaje certero.

Cuando se habla de usuarios se considera a entidades gubernamentales, educacionales, centros de investigación y todo tipo de ente que se encuentre dedicado a proyectar conclusiones que puedan ser de utilidad en el proceso de toma de decisiones de índole poblacional.

El significado de muchos de los términos utilizados en el área en la cual se aplica este trabajo son explicados en detalle en los siguientes capítulos, el capítulo III a lo dedicamos a la definición de términos como el mencionado anteriormente, dando principal importancia a la población y las características necesarias para cuantificarla, en el capítulo II damos una síntesis de SIG, este capítulo es la base de esta investigación, incursionamos en esta área buscando formas de representar geográficamente las áreas censadas, y las herramientas que actualmente se utilizan para manipular la información geográfica que se deriva de un censo.

En estos capítulos teóricos se consideraron definiciones que fueron relevadas de diversos especialistas en el área y se recopilaron además diversos artículos que fueron la base para emprender este trabajo. La Arquitectura resultante no es más que un fiel reflejo de lo relevado.

Se tomo como principal fuente de datos al INDEC y se analizaron las estadísticas disponibles, ya que no es posible acceder a datos reales sino sólo a los resultados de los censos.

Se relevaron además encuestas censales las cuales sirvieron de base para poder caracterizar a las entidades del modelo resultante.

El cuarto capítulo esta destinado al problema en si [24], el mismo se origina cuando se calculan Indicadores demográficos sobre regiones geográficas pequeñas. Describiremos su aplicación y detallaremos las características que tomaremos en cuenta para el diseño de nuestro modelo, al final de este capítulo fundamentaremos el uso del paradigma de objetos y del lenguaje de modelado UML que se utilizará para el análisis, junto con él investigamos herramientas que lo apliquen, utilizamos Poseydon Community Edition, la misma puede ser descargada de Internet. [29]

La fundamentación se basa en lo leído de expertos de área que explican la utilización del paradigma para desarrollar sistemas basados en SIG, el paradigma sobre el cual no teníamos experiencia previa ya que actualmente los proyectos en los que participamos son netamente relacionales, fue en este punto donde encontramos la mayor dificultad, en la representación.

El último capítulo cierra este trabajo presentando todo el análisis y la documentación sobre lo relevado, mostramos el modelo conceptual a partir de todo lo investigado en el ámbito de censos y el problema que analizamos en el Capítulo IV, luego describimos el modelo geográfico basándonos en una arquitectura geográfica la cual analizamos [30][22] y en lo que aprendimos de SIG en el Capítulo II, la arquitectura resultante muestra la unión de ambos modelos agregando la funcionalidad necesaria para que interactúe con el usuario.

Para finalizar damos las conclusiones del trabajo, y con ellas el cierre de una etapa académica que con altibajos logramos superar, el escenario, acerca del cual surgieron las ideas que le dan vida a este trabajo, se relacionan con todo el conocimiento adquirido a lo largo de la carrera, los avances tecnológicos a los que nos enfrentamos en nuestro desempeño laboral, entre ellos la evolución de la Ingeniería de Software así como la aparición de un nuevo paradigma que nos lleva a desarrollos alejados de lo que aplicamos en el día a día.

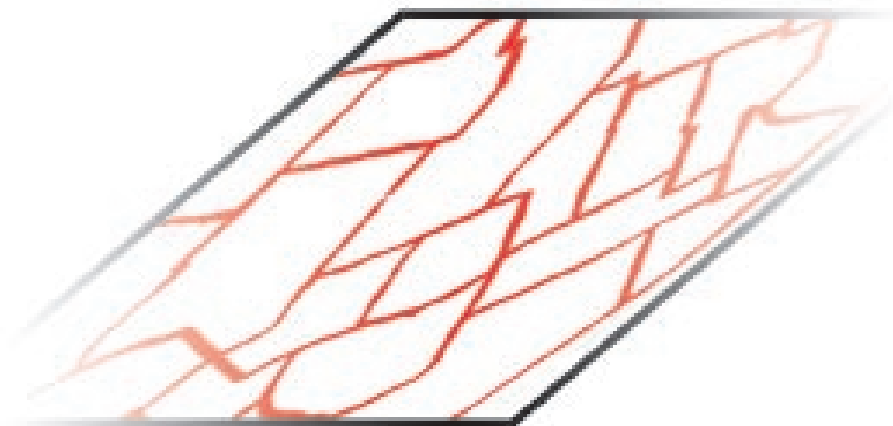
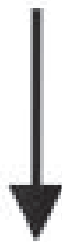
Impulsaron la investigación haber encontrado un problema real que aplique conceptos

que no habíamos explorado con profundidad, poder poner en práctica la formación que recibimos en cada una de las asignaturas en las que fuimos evaluadas, fue una buena experiencia sobre la que sentimos que pudimos demostrar que se nos ha formado para resolver problemas, seleccionando la mejor alternativa tecnológica existente, con o sin conocimientos previos.



## CAPÍTULO II

### SISTEMAS DE INFORMACIÓN GEOGRÁFICA



## INTRODUCCIÓN

Como punto de partida de nuestra investigación, hemos centrado nuestra atención en los sistemas de información geográfica (SIG), investigamos acerca de sus características descubriendo así la diversidad de campos en los que pueden ser utilizados.

Una de las características más importantes de un SIG, es que, al igual que los sistemas de información de datos, estos sistemas almacenan y manipulan información, en este caso geográfica.

Entonces podemos decir que un SIG también es un sistema de información utilizado como una herramienta para la toma de decisiones, responsable de mantener la calidad y la exactitud de la información que administra.

Hablamos de la información en general, debido a que un SIG puede ser utilizado para múltiples propósitos en diversas disciplinas que necesiten contar con información geográfica de cualquier índole.

En este capítulo damos una serie de definiciones de SIG y de factores que están relacionados con información geográfica. Mencionamos también una serie de características relevantes según el tipo de datos y las formas de almacenar este tipo de información. Para mostrar su aplicación damos ejemplos que nos fueron útiles para entender los beneficios de un SIG.

Un SIG gestiona una base de datos espacial, permite la creación y estructuración de los datos partiendo de fuentes de información como los mapas, la teledetección, bases de datos existentes, etc. Además de esto, posibilita el análisis, visualización y edición de los datos de la base de datos, un SIG cuenta con herramientas que permiten crear nuevos datos derivados de los existentes. [1]

De esta definición se desprende que un SIG es un sistema de información que permite manipular datos geográficos. Esto significa que los datos geográficos tienen coordenadas geográficas reales asociadas, permitiendo manejar y hacer análisis con datos reales, como longitudes, perímetros o regiones.

## MAPAS

Un mapa es una colección de conceptos almacenados, se considera también como una clase de sistema de información, la información derivada de él es usada para la toma de decisiones.[5]

Ya que los SIG hacen referencia a la información geográfica, los mapas resultan fundamentales, éstos representan características como por ejemplo vegetación, clima y ubicación entre otras. Con respecto a la ubicación podemos decir que permite determinar

la posición de un punto en coordenadas, describiendo así la proximidad y dirección de los objetos, y su conectividad.

Pero si nuestro fin es, por ejemplo, representar el clima de una región, tener tantas características de más puede hacer más complejo el análisis, es por esto que los mapas se pueden clasificar según la información que contengan. Por ejemplo si lo que se desea es conocer la cantidad de enfermos que existen en una región para algún tipo de enfermedad en particular, el mapa solo deberá representar de alguna forma dicha cantidad, sin agregar características que no son relevantes en ese contexto.

## BASE DE DATOS

Conocida ampliamente en nuestra área de trabajo, podemos decir, en base a lo que leímos hasta ahora, que es la fuente de datos principal de un SIG, sin ella no tiene propósito alguno tener un mapa, el cual terminaría siendo solo un dibujo, perdiendo toda importancia y dejando de ser un sistema de información.

En general una base de datos puede ser cualquier recopilación de datos con sentido práctico, como bien sabemos su organización es parte fundamental en cualquier sistema, y en particular encontramos que es la razón de ser de un SIG.

Cuando hablamos de sentido práctico nos referimos a que la información que sea almacenada en ella debe ser la necesaria, esto es un punto clave en un SIG, ya que información de más, genera un mayor costo y dificulta el acceso.

Los aspectos que se deben tener en cuenta sobre la información que ingresemos a nuestra base de datos es que debe ser temporalmente precisa y completa, de no ser así los datos que devuelva el SIG no serán fiables.

Hasta ahora sabemos que las fuentes de información más importantes de un SIG son generalmente mapas y bases de datos, ahora analizaremos los componentes de un SIG y su contexto.

## COMPONENTES

Un SIG generalmente puede ser caracterizado a través de cuatro componentes:[1]

### Entrada de datos:

Cuando seleccionamos las fuentes de información, estas deben ser convertidas al formato que utilice el SIG. Este procedimiento de conversión puede ser simple o complejo dependiendo de los formatos de los datos de entrada.

La construcción de la base de datos puede tener un costo entre cinco a diez veces más

que el del software y hardware que se utilice, dado que la entrada inicial de datos puede llevar desde meses hasta años. Es por esto que esta es la etapa más delicada, debido a que es muy costoso corregir errores en datos ya ingresados.

#### Gerenciamiento de datos:

este componente incluye las funciones necesarias para almacenar y devolver datos de la base de datos. Los métodos usados para implementar estas funciones afectan directamente la performance de las operaciones que se realicen con los datos. Es por este motivo que cobra importancia la manera en que los datos se encuentran estructurados dentro de la base, las relaciones entre ellos y las restricciones que existen entre estas relaciones, ya que estos factores son los que determinan la velocidad de la extracción de los datos cuando se ejecuta una consulta. Por ello resulta necesaria la intervención de un experto durante el diseño y análisis de los procedimientos a implementar en las bases de datos SIG.

#### Manipulación y Análisis de la Información:

Las funciones de manipulación y análisis son las que determinan la información que puede ser generada a través del SIG. Para especificar estas funciones necesitamos involucrar a los usuarios a fin de poder anticipar los datos que se analizarán, así como también para determinar cual debe ser el nivel de performance en cada una de las funciones solicitadas.

#### Generación de Resultados:

Se refiere a los reportes que pueden ser generados, mostrando los resultados de las consultas que se procesan en un SIG. Estos reportes pueden tener formato de mapas, tablas de datos o texto plano. Dicho formato se define en base a los requerimientos de los usuarios, de manera que es importante su participación a la hora de elaborar la especificación.

## CONTEXTO

También vimos que las características de un SIG varían de acuerdo al contexto en el que resida, por ejemplo un censo nacional va a usar un SIG para asistir en sus investigaciones y puede también proveer servicios de análisis a otros departamentos o al público en general. Al mismo tiempo los militares pueden desarrollar un SIG aún más detallado de la misma área geográfica estando su uso rigurosamente restringido y controlado.

Existen otros sistemas, como CAD (Computer Aided Design and Drafting) y DBMS (Data Base Management System) que se utilizan actualmente para la manipulación de datos georreferenciados y que pueden ejecutar muchas de las mismas operaciones que un SIG, pero un SIG provee además la habilidad de integrar datos georreferenciados. Esto



implica que con un SIG podemos realizar búsquedas y superposiciones espaciales y crear información al integrar los datos originales de diferentes maneras y desde distintas perspectivas.

## PROCESO DE ABSTRACCIÓN

Por lo explicado hasta ahora, lo que estamos haciendo al crear un SIG, es tomar entidades del mundo real, de las que necesitamos generar conclusiones con algún propósito, abstrayendonos de la realidad. En [1] se describe este proceso como se muestra en la figura 2.1.

En ella se puede ver que el proceso comienza al referenciar entidades del Mundo Real, recolectando solamente la información necesaria, organizandola para facilitar su análisis y la posterior toma de decisiones, que luego se implementaran en el Mundo Real

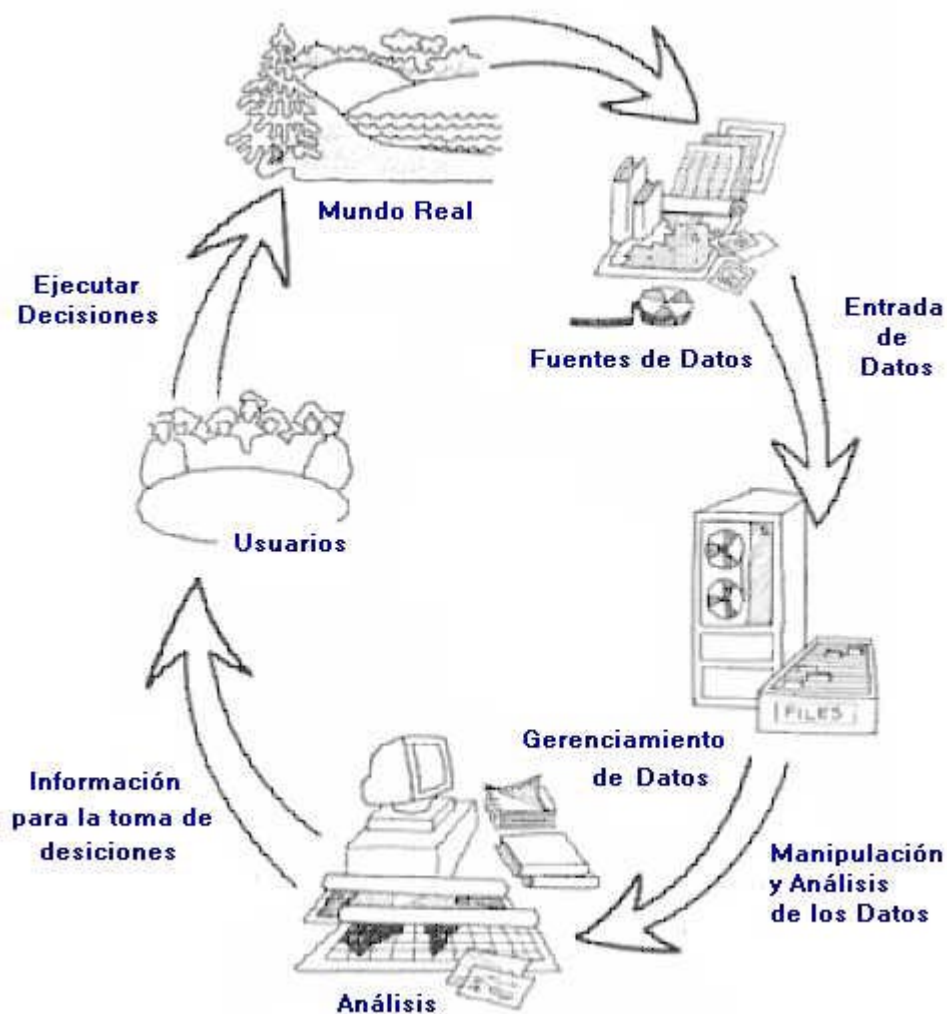


Figura 2.1 La información geográfica empieza y termina en el mundo real.



La información geográfica a tener en cuenta sobre la entidad del mundo real a representar describe un fenómeno del mundo real en una ubicación determinada y para un momento determinado. Además, describe las dimensiones físicas de la entidad.

Cuando mencionamos dimensión física nos referimos por ejemplo a la cantidad de personas en una ciudad, al ancho de una calle, etc. Por otro lado la ubicación se especifica a través de un sistema de coordenadas como pueden ser latitud y longitud.

Este proceso de asignar coordenadas geográficas a un objeto o estructura se denomina **Georreferenciar**, al realizarlo hacemos referencia a un sistema mediante el cual, es posible estimar las coordenadas actuales de un objeto con respecto a una posición fija en la superficie terrestre. Este concepto caracteriza a un SIG, pues este manipula datos georreferenciados. Estos datos también reciben el nombre de datos espaciales. [1]

$$\text{Datos georreferenciados} = \text{Datos espaciales}$$

La georreferenciación puede utilizar un sistema de coordenadas establecido para un determinado sistema de proyección, o representar los objetos mediante una clave o índice, generalmente con significado administrativo como puede ser dirección, código postal, etc., esto permite identificar una posición aproximada.

En cuanto a referenciar características a través de un índice, podemos decir que obviamente no tiene la misma precisión que la obtenida cuando se utiliza un sistema de coordenadas. Es útil en realidad cuando queremos aprovechar de forma inmediata la información disponible georreferenciada a través de un sistema de coordenadas, ya que la información estaría directamente vinculada al índice o clave definido.

## SISTEMA DE PROYECCIÓN

Normalmente los SIG proporcionan funcionalidades para realizar cambios de coordenadas entre varios sistemas. Los sistemas de proyección están pensados para resolver el problema de proyectar la superficie curva de la tierra en un sistema plano. Aunque todo sistema de proyección distorsiona la realidad, podemos mantener sin distorsión el área utilizando proyecciones equivalentes, distancias equidistantes o ángulos conformes. El sistema de proyección es importante porque cuando se necesiten comparar regiones estas deberán tener proyecciones equivalentes para poder trabajar con ellas.

Uno de los sistemas de proyección más utilizado es UTM (Universal Transversal Mercator), que se obtiene proyectando las coordenadas sobre un cilindro cuya directriz es un meridiano terrestre con distorsión nula. En este caso, la georreferencia se expresa mediante un identificador de zona y dos coordenadas latitud y longitud en metros, según los ejes Este-Oeste y Norte-Sur respectivamente.

Otra forma de proyectar son las coordenadas planas, en las que no se tiene en cuenta la curvatura terrestre, en aquellos casos en los que las áreas de interés sean las dimensiones moderadas, ya que en otro caso se producirán inconsistencias por las distorsiones introducidas.

## TIPOS DE DATOS

Venimos mencionando que la información que queremos representar al utilizar un SIG consiste en datos geográficos y no geográficos, para hacerlo se necesita conocer la manera en que están almacenados para poder integrarlos. Esta es la razón por la que necesitamos definir dos conceptos que son: datos espaciales y datos alfanuméricos. Estos son los tipos de datos que manipula un SIG para poder representar las características antes mencionadas.

Encontramos que los datos espaciales y los alfanuméricos poseen características específicas y diferentes requisitos para su almacenamiento, procesamiento y representación. Por esta razón los explicaremos por separado y al finalizar daremos las aplicaciones y beneficios que nos brinda un SIG al momento de manipular estos dos tipos de datos de manera integrada.

### Datos Espaciales

Los datos espaciales son descripciones digitales de las entidades del plano. Suelen incluir las coordenadas, reglas y símbolos que definen los elementos cartográficos en un mapa.

#### *Características*

Los datos espaciales tienen ciertas características que es necesario tener en cuenta a la hora de representarlos, estas son las que nos van a ayudar a identificar las relaciones existentes entre ellos, nos referimos a:

#### Relaciones topológicas:

Hacen referencia a la ubicación relativa de dos o más entidades, como pueden ser alguna posición relativa de dos casas. Estas relaciones pueden estar en los datos o ser deducidas a partir de la proximidad, solapamiento, etc.

#### Clasificación:

Los objetos pueden ser clasificados en distintas clases o categorías, por ejemplo, la capa de transporte que comprende autopistas, carreteras, etc.

### Agregación:

Los objetos del mundo real pueden ser definidos como composición o agregación de otros objetos, por ejemplo un colegio se puede considerar como la agregación de edificios, campos de juego, etc.

### Asociación:

Podemos hacer una asociación entre un edificio y la calle más cercana.

## REPRESENTACIÓN DE LOS DATOS ESPACIALES

La información geográfica puede representarse de diferentes formas: geométricas ó cartográficas, mediante puntos, líneas, polígonos ó como imágenes por medio de fotografías ó como atributos como por ejemplo una dirección. Mas adelante explicaremos los formatos más usados actualmente.

### Formatos de Representación

#### *Raster*

Es un método para almacenar, procesar y mostrar los datos espaciales. En este formato cada área se divide en columnas y filas, las cuales forman una matriz en la que cada celda debe ser rectangular o cuadrada. Cada celda de esta matriz contiene la ubicación como un atributo, denotado por sus coordenadas, además cada línea se define por todos sus puntos intermedios. Cada punto es denominado pixel o “bit”. El problema que presenta este formato es que pierde resolución, los límites de estructuras poligonales no se pueden identificar correctamente. Por esto hay que tener ciertas consideraciones a la hora de definir el tamaño de las celdas. El tamaño de cada píxel debe ser la mitad de la distancia más pequeña a representar [2].

#### *Modelo Raster*

Entonces, basándonos en la explicación anterior vamos a definir qué es lo que debe tenerse en cuenta para trabajar con un modelo que utilice este tipo de formato para representar datos espaciales.

Como mencionamos en la definición, un Modelo Raster debe representarse en pequeños rectángulos o cuadrados cuyo tamaño será fundamental ya que como sabemos, el tamaño determinará su resolución.

Estos rectángulos se representan por medio de un píxel, en matrices de puntos de forma cuadrada o rectangular que contienen valores numéricos y que representan las entidades cartográficas y sus atributos a la vez. Los modelos lógicos de poca complejidad son

basados en el modelo conceptual raster, ya que la georreferenciación y la topología están implícitas por las posiciones de columna y fila, de cada píxel en la matriz. [3]

Existen varias técnicas [1] para crear modelos raster, entre ellas están los métodos *quadtree*, *octree* y *run-length encoding* que son los más utilizados.

#### Quadtree - Octree

Ambos métodos consisten en dividir un mapa en una estructura jerárquica basada en la descomposición recursiva del espacio en cuadrantes y octantes respectivamente, resultando una estructura de árbol, logrando reducir el espacio de almacenamiento y el tiempo de procesamiento de los datos que demanda el formato raster.

Conceptualmente la construcción de un quadtree puede ser tomada como un proceso que subdivide un mapa recursivamente hasta que la porción de mapa a dividir contenga el mismo tipo de característica como por ejemplo el área un bosque, una zona residencial, comercial, etc.

Recordemos que en el formato raster convencional, el mapa se representa por un largo número de pequeñas celdas con el mismo valor, de esta manera este método reduce el almacenamiento. En [1] se describe el método quadtree como se muestra en la figura 2.2.



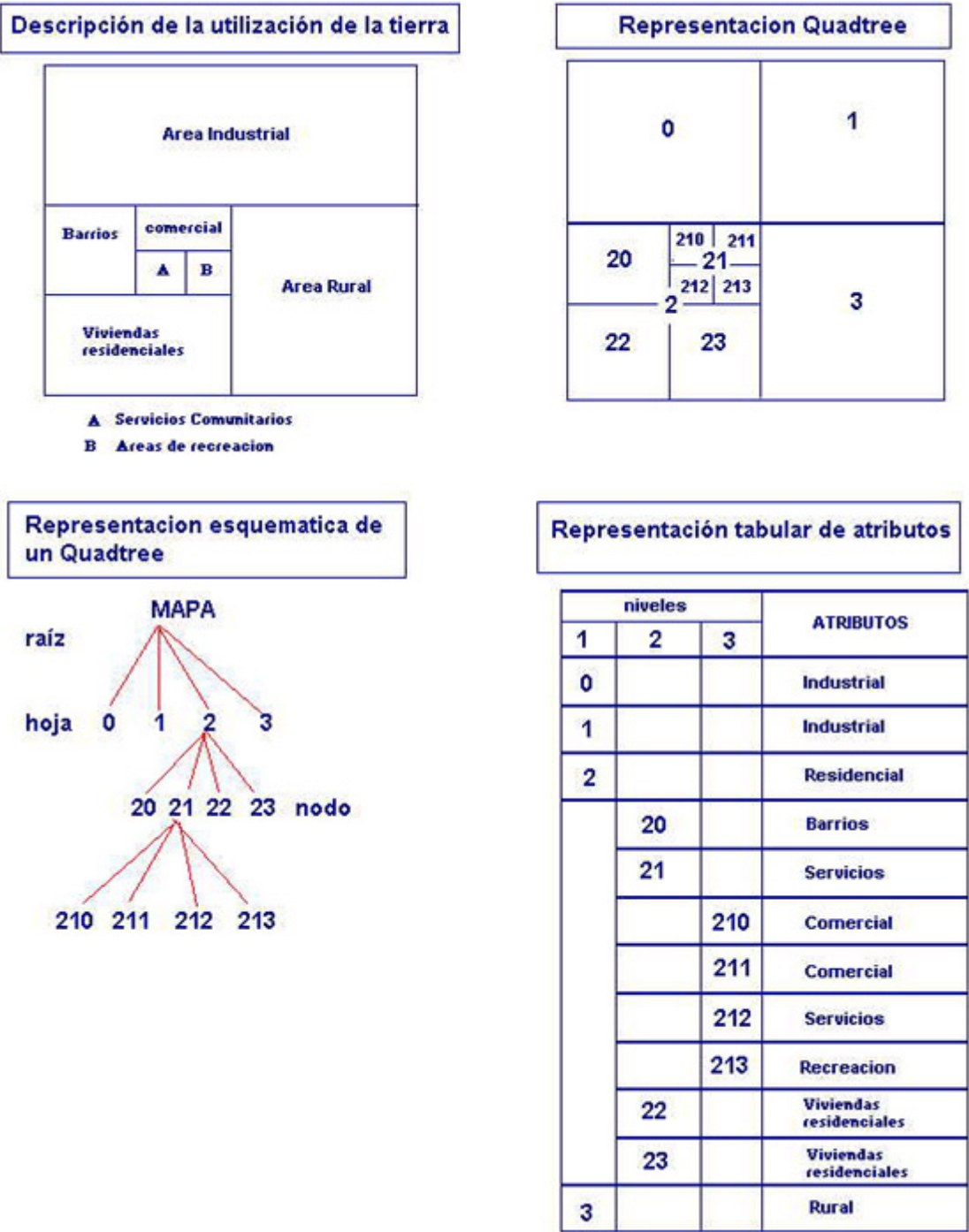


Figura 2.2 Representación Quadtree.

Run-Length Encoding

Por otro lado, la técnica *Run-length encoding* se basa en que los objetos frecuentemente se extienden sobre áreas mayores que un único píxel, así, este método, en lugar de guardar los valores de cada uno de los píxeles, agrupa las filas de una matriz raster en bloques con idéntico valor como se muestra en la figura 2.3. Es decir, codifica los valores de los píxeles agrupándolos en el par ordenado (run-length, valor).

Por ejemplo, si los valores de una fila de píxeles que representan una imagen en blanco

y negro fueran 000011100, usando este método se almacenarán 403120

Donde (40)3120

(Contador, valor del dato)

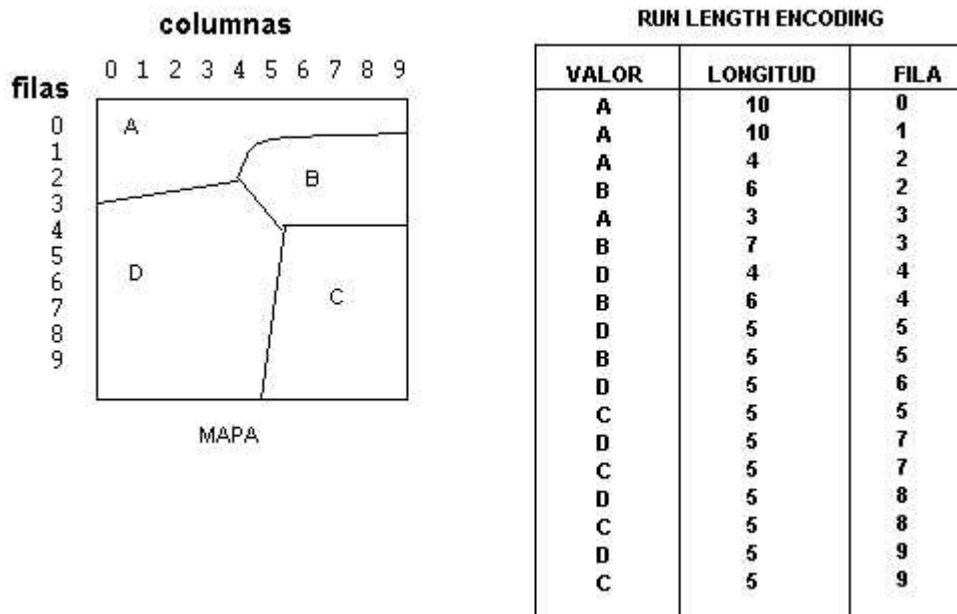


Figura 2.3 Método Run-Length Encoding.[1]

Mencionamos que los métodos anteriores logran reducir el espacio de almacenamiento y en consecuencia el tiempo de procesamiento. Estos son los principales inconvenientes que se presentan cuando se utiliza el modelo raster, debido a que tiene una organización muy simple de datos, lo cual permite realizar con gran facilidad ciertos procesos de análisis, como por ejemplo la superposición de planos, esta operación es muy sencilla de programar mediante operaciones con matrices, resultando muy costosa cuando la representación de los datos está en otro formato.

Si bien los datos almacenados bajo un modelo Raster pueden ser fácilmente imprimibles y comparables con otro gráfico del mismo modelo, los inconvenientes que presenta son el volumen de almacenamiento, la baja representación espacial y la alta complejidad para realizar el análisis de datos espaciales.

La precisión cuando se referencien entidades espaciales con este modelo se encuentra limitada por la porción del territorio que representa un píxel. Además hay que tener en cuenta que muchas veces no se especifica como está georreferenciada la celda, si es respecto a su ángulo superior izquierdo o a su ángulo inferior izquierdo o respecto a su centro.

Mediante los métodos run.length encoding y quadtree se busca minimizar el almacenamiento, pero aunque se logre, otra consideración que hay que tener presente a la

hora de elegir el modelo Raster, es que no representa explícitamente la topología. Por lo tanto en las aplicaciones en que sea esencial su empleo, este modelo no podrá ser utilizado.

### *Vectorial*

Este formato trabaja con unidades homogéneas, puntos, líneas y polígonos, comparado con el formato Raster, las unidades son relativamente menores en número y variables en tamaño. En un archivo generado con este formato, la cantidad de elementos puede llegar a las decenas de miles, pero no a los millones como ocurre comúnmente en un archivo con formato Raster. La posición de estas unidades homogéneas está definida usando un determinado sistema de coordenadas. [1]

### *Modelo Vectorial*

Este modelo proporciona información precisa acerca del posicionamiento en el espacio de las entidades del Mundo Real a representar. Para esto, utiliza las tres primitivas básicas mencionadas en el punto anterior. [6]

#### Nodos (o punto):

Específica una entidad geográfica que puede ser localizada dentro de un sistema cartesiano, haciendo referencia a un par de coordenadas.

Se utiliza el punto para representar un fenómeno geográfico, un lugar, como puede ser una plaza dentro de un plano, y que es muy pequeño como para representarlo en función de líneas o polígonos.

#### Líneas (o arcos):

Son objetos definidos por un nodo inicial y un nodo final. Se utilizan para representar por ejemplo una calle, un río.

#### Polígonos (o áreas):

Son objetos formados por una serie de líneas que limitan un área continua de dos dimensiones. Con ella se podría representar los límites una ciudad.

Entre los elementos que conforman este modelo, existe una serie de relaciones tales como que una línea se define por más de dos puntos, solo si es necesario representar alguna entidad geográfica en su trayecto, o que un polígono está limitado por una serie de líneas, lo cual constituye una mínima definición topológica, tal y como mostramos en la figura 2.4.

Normalmente se almacenan relaciones del tipo: punto origen, punto final de línea y



relación ordenada de los puntos internos. Las secuencias ordenadas de las líneas, definirán entonces un polígono.

El detalle con el que se almacenan las relaciones es un compromiso entre la eficiencia del proceso y la precisión con la que se modela la realidad.

Los procedimientos de análisis en este modelo son más laboriosos, pero más precisos que en el modelo raster, ya que implican la resolución analítica de intersecciones entre arcos, la determinación de áreas y la evaluación de posiciones relativas entre elementos diferentes (punto / polígono, punto / línea, etc.).

La posición de los datos puede ser georreferenciada directamente, por medio de un sistema de coordenadas matemáticamente exacto, o indirectamente, utilizando por ejemplo la dirección postal. En ambos casos la solución es muy eficaz. Los atributos no espaciales son almacenados en una base de datos alfanumérica (no espacial) interrelacionada con una base de datos cartográfica (espacial), ofreciendo con ello posibilidades muy distintas a las que ofrece el modelo Raster.

Todas las características topológicas se explicitan hasta el último detalle, en la figura 2.4 mostramos un ejemplo de esta característica. De hecho, la resolución de este formato sólo se ve limitada por el número de bits usado para representar un valor en la computadora que se este utilizando. Por lo tanto su resolución es mucho mejor que los tamaños de las celdas utilizadas en el formato Raster.

Este modelo es parecido a la percepción humana del espacio y por lo tanto, está más de acuerdo con la cartografía tradicional por lo que resulta más intuitivo. Podríamos, por ejemplo, representar las limitaciones de una ciudad con un polígono, una plaza con un punto y un río con una línea.

La principal ventaja de este modelo respecto del modelo raster es su capacidad para expresar las relaciones espaciales existentes entre las entidades, esto es, la información topológica, que es la que dota al modelo de la semántica necesaria para representar el conocimiento territorial.



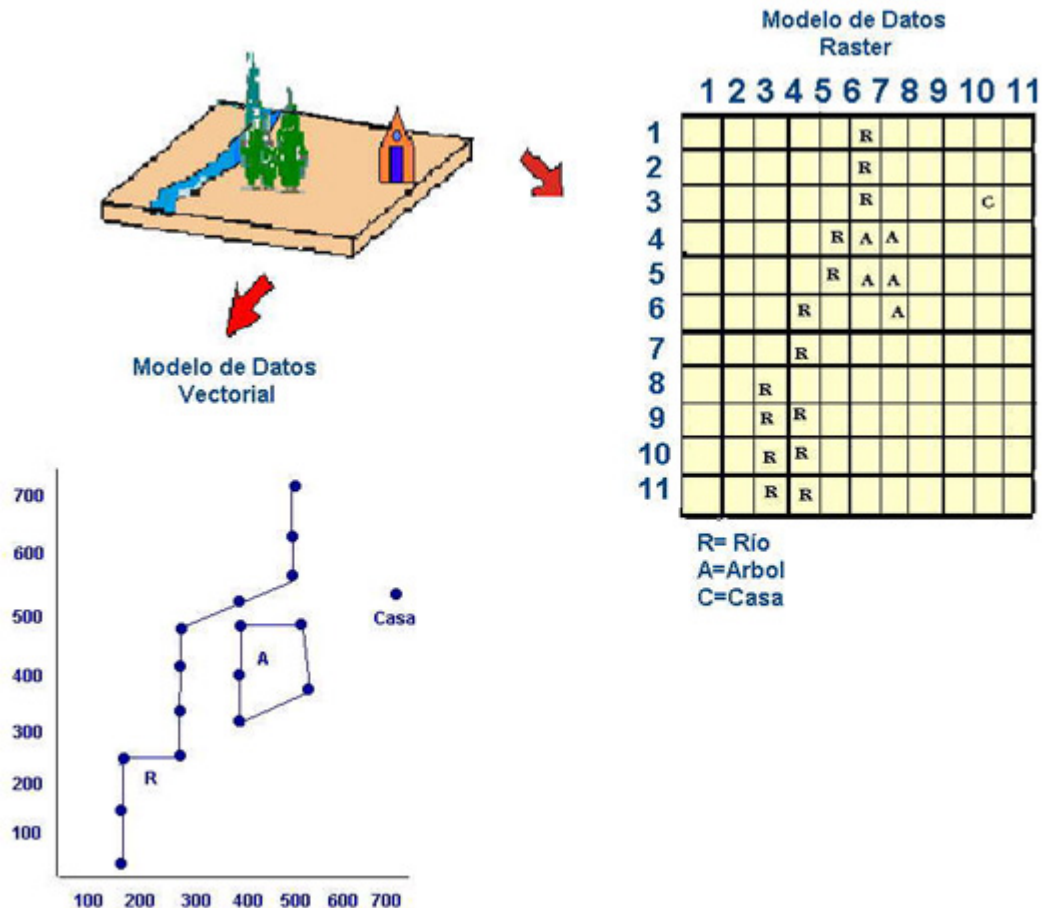


Figura 2.4 Comparación del modelo Raster y Vectorial. El Mundo Real es representado y muestra claramente como el Modelo Vectorial captura la geografía en su representación.[1]

## Condiciones a Tener en Cuenta

A pesar de todas las diferencias planteadas en los modelos vectoriales y raster, actualmente los principales sistemas SIG combinan ambos tipos de estructuras. El debate raster/vectorial ha evolucionado de la cuestión “¿cual es mejor?” A la cuestión “¿bajo qué condiciones es mejor uno que otro, y cómo podemos combinarlos de manera flexible?”

Por esto, a la hora de decidirse por un formato se debe tener en cuenta:

- Usar el formato vectorial para la realización de gráficos y mapas precisos.
- Usar el formato vectorial para análisis de redes (cableados eléctricos y telefónicos, rutas de transporte, etc.).
- Para la superposición y combinación de planos es más rápido y barato el modelo raster.
- Usar el método raster cuando se trabaja con representaciones y simulaciones de superficies.

- Utilizar el formato raster y vectorial en combinación cuando es necesario representar líneas con precisión (vectorial) y superficies rellenas (raster)
- Disponer de algoritmos de conversión de vectorial a raster y viceversa.
- Recordar que se pueden editar simultáneamente datos raster y vectoriales.

## Datos No Espaciales o Alfanuméricos

Cuando se mencionan datos alfanuméricos en un SIG nos referimos a las características que poseen las entidades espaciales, estas características son almacenadas en formatos comunes para este tipo de información, es decir, como en cualquier base de datos. Se refiere por ejemplo a: nombre, código postal, etc. En una base de datos de este tipo podemos encontrar los siguientes tipos de datos alfanuméricos:

### Atributos Alfanuméricos:

Describen las características de las entidades espaciales. Se relacionan con dichas entidades a través de identificadores comunes que se almacenan tanto en el registro alfanumérico como en el gráfico. Un sistema SIG debe ser capaz de realizar consultas o análisis sobre los atributos alfanuméricos de forma independiente y generar mapas basados en dichos atributos.

### Datos Geográficamente Referenciados:

Este tipo de datos describen incidentes o fenómenos que se producen en una localización específica. A diferencia de los atributos, estos datos no describen una entidad espacial sino que proporcionan información (número de edificios permitidos en una zona, número de accidentes en un cruce, inspecciones de salud en un barrio, etc.) asociada a una localización geográfica. Este tipo de datos se almacena y gestiona de forma separada y no se relaciona directamente con las entidades geográficas de la base de datos del SIG.

Los mecanismos que existen para optimizar el acceso a este tipo de información son:

### Indices Geográficos:

Los índices geográficos se utilizan en un SIG para seleccionar, relacionar y recuperar datos en función de su localización geográfica, de forma similar a como actúan los índices en una base de datos tradicional; no constituyen información en sí y únicamente sirven para mejorar los accesos.

### Relaciones Espaciales:

Proporcionan información sobre las relaciones entre los distintos tipos de entidades espaciales y objetos geográficos, como son las existentes entre las líneas o las adyacencias entre los polígonos.

## Aplicaciones de SIG

Nos resulta interesante rescatar algunas de las aplicaciones SIG mas usadas. Ellas se aplican a distintos sectores, tanto públicos como privados, ya que representan una herramienta importante de ayuda en el momento de obtener información geográfica. Entre las aplicaciones consideramos importante mencionar [4]:

- Aplicaciones de cartografía automatizada, están relacionadas con sectores públicos y privados que se encargan de obtener y mantener planos digitales de distintas áreas con el fin de venderlos a otros sectores a los que les puede interesar con diversos fines.
- Una utilidad interesante es la aplicación de los SIG para almacenar y mantener datos referentes a redes eléctricas, gas, cloacas, etc. Para las empresas encargadas de mantener este tipo de información es de mucha utilidad este tipo de sistemas, ya que les permite relacionar información alfanumérica, como puede ser número de medidor, con sus unidades geográficas.
- Desde el punto de vista urbano, una utilidad importante es la posibilidad de obtener información catastral y urbanística. Podemos generar esta información mediante la utilización, en forma conjunta, de los métodos Raster y Vectorial, superponiendo distintas representaciones de un área particular. Esto nos permite obtener, por ejemplo, información en forma automática de las zonas catastrales que han realizados cambios de infraestructura. Se puede, por ejemplo, determinar fácilmente, mediante superposición de cartografías, si en una parcela se ha realizado alguna modificación inmobiliaria, como puede ser una pileta.
- Desde el punto de vista del impacto en el medio ambiente, los SIG facilitan la evaluación para la ejecución de proyectos en los que estén involucrados, por ejemplo, la utilización de productos contaminantes, en cuyo caso nos permite tomar medidas correctoras, como puede ser la reforestación.
- Desde el punto de vista social, nos permite obtener información sobre centros sanitarios, escuelas, etc. que existen en una zona dada, y nos ayudan a planificar la ubicación estratégica de centros nuevos. Así como también nos permite realizar estudios epidemiológicos relacionando la incidencia de enfermedades con el entorno vital.
- Desde el punto de vista Urbano, podemos usar el SIG para modelar el comportamiento del tráfico estableciendo modelos de circulación por una calle en función de las condiciones de tráfico y longitud. Asignando un costo a los puntos en los que existe un semáforo, se puede obtener información muy útil, como puede ser: deducir el camino más corto en distancia o en tiempo entre dos puntos.

## Beneficios

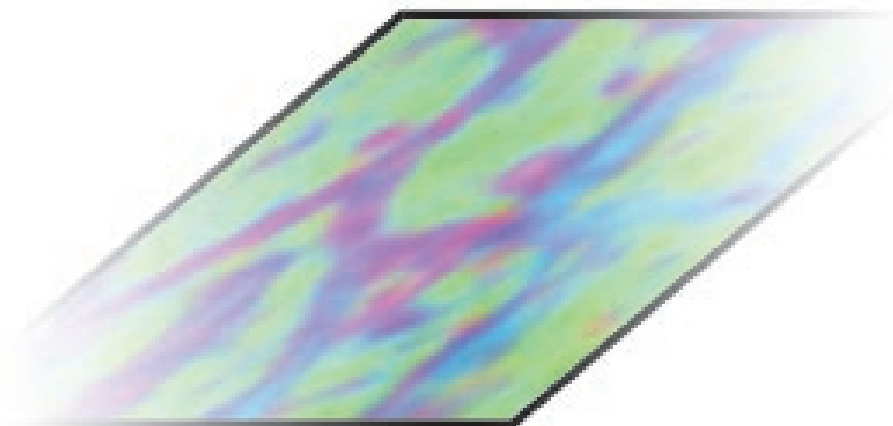
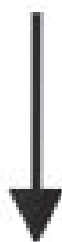
También hemos investigado acerca de los beneficios económicos que puede brindar un SIG a una empresa. Entre ellos destacamos los siguientes:

- Ahorro de tiempo en producción, mantenimiento y administración de mapas.
- Información exacta, actualizada y centralizada.
- Acceso rápido a los datos.
- Reducción de actividades redundantes o tediosas.
- Análisis complejos imposibles de hacer por métodos tradicionales.
- Menores costes de operación.
- Ayuda a la toma de decisiones para la realización de inversiones más efectivas. La calidad e las decisiones tomadas depende de la calidad de los datos ingresados y el modelo del espacio usado en el análisis.
- Intercambio o venta de información impresa o en soporte magnético.
- Creación de nuevos servicios, derechos por el uso de las bases de datos, etc.
- Obtención inmediata de estadísticas, mapas temáticos, etc.
- Mejora del servicio a los clientes.
- Fácil acceso a la información (por dirección, calle, número de parcela, etc.).
- Análisis e informes de gran calidad (mapas temáticos, estadísticas, listados, etc.).
- Eliminación de información redundante en distintos departamentos, al estar totalmente integrada.
- Incremento de la productividad.



# CAPÍTULO III

## CENSOS POBLACIONALES INDICADORES DEMOGRÁFICOS



## INTRODUCCIÓN

En este capítulo explicaremos los conceptos sobre los que necesitamos profundizar acerca de la información censal.

La información censal se utiliza principalmente para el análisis, seguimiento, revisión y formulación de políticas públicas de desarrollo; así como de los programas, proyectos de inversión física y actividades programadas anualmente, para el mejoramiento del nivel de vida de la población, fundamentalmente de la más necesitada, a través del conocimiento de las necesidades y carencias básicas de bienes y servicios.

Desde el punto de vista de la administración pública, la información sobre aspectos demográficos, sociales y económicos determina la política pública y condiciona el tamaño y las funciones del Gobierno, ya que a más población, más demanda de servicios y más cambios en la distribución y estructura del gobierno. También condicionan el tamaño del Organismo Legislativo, el cambio de categoría de los lugares poblados, la creación de municipios y el presupuesto que el gobierno debe asignar a las municipalidades.

Entonces lo que motiva a los usuarios a demandar información censal es la necesidad de ampliar los conocimientos que tienen sobre el desarrollo de alguna población en particular, lo cual les permite conocer las consecuencias de los movimientos migratorios, obtener información sobre las condiciones de producción económica, pudiendo determinar cuando se producen cambios socioeconómicos.

Esta demanda de información genera una investigación estadística para obtener datos censales. Para realizar una investigación estadística lo primero que debemos hacer es definir el objeto a estudiar, luego debemos diseñar indicadores del fenómeno que se quiere medir y finalmente debemos efectuar las tareas de obtención de datos de la población, para lo que debemos tener en cuenta la implementación de tareas de validación a la hora de procesar los datos y finalmente obtener la información estadística que debemos entregar a los usuarios que la solicitaron, con el propósito de que estos realicen el análisis e interpretación de la información, tomándola en cuenta en su toma de decisiones, introduciendo de esta manera nuevas consecuencias socioeconómicas, las cuales se transforman en otras demandas de información, generando nuevas investigaciones estadísticas.

Para facilitar la comparación de los resultados, los censos se realizan periódicamente, en ellos se relevan características básicas sobre población y vivienda, actividad económica y agropecuaria. Este relevamiento se realiza a través de encuestas las cuales muestran los resultados de una muestra a la población total. Esta metodología se aplica en la Encuesta Permanente de Hogares; la Encuesta Nacional de Gastos de los Hogares; la Encuesta Industrial Mensual y Anual; la Encuesta Nacional Agropecuaria, etc.

Los índices e indicadores que debemos producir los define un usuario, los mismos se generan en base a los datos primarios o secundarios que se generan de las encuestas realizadas, más adelante explicaremos como esta información permite efectuar un seguimiento preciso y sintético sobre la evolución de distintos fenómenos.

## OBJETIVOS DE UN CENSO

- Dotar al país de información básica, confiable y oportuna sobre la población, locales de habitación, servicios y la estructura agraria del país, lo cual permite evaluar y mejorar los planes, programas y proyectos de desarrollo urbano y rural.
- Mejorar las condiciones de la producción estadística en el sistema estadístico nacional, teniendo como marco de referencia los resultados de los censos nacionales integrados.

### Datos estadísticos necesarios

- Información sobre el tamaño, composición y distribución de la población y de sus características sociales y económicas.
- Información sobre los locales de habitación, sus características principales y las condiciones habitacionales de los hogares.
- Conocer información sobre la tenencia y uso de la tierra, las características principales de las unidades de producción agropecuaria, tipo y diversidad de los productos, origen y destino, volumen de producción, áreas de producción agropecuaria y la del medio ambiente.
- Material cartográfico actualizado de todo el territorio nacional.
- Resultados censales, un marco muestral que facilite el diseño y ejecución de futuras investigaciones estadísticas.

### Cómo Podemos Utilizar la Información de los Censos

Como mencionamos anteriormente los datos censales son importantes, tanto para el sector público como privado. Esto ocurre debido a que posee características de universalidad, simultaneidad y periodicidad, siendo útil para la ejecución de proyectos y elaboración de estudios técnicos sobre condiciones de vida, mortalidad, fecundidad, migración interna e internacional, características socioeconómicas de la población activa, necesidades de vivienda, destino de la producción, usos principales de la tierra, volumen y extensiones de la producción agropecuaria y forestal, así como las condiciones ambientales y de los recursos naturales, y las del mercado.

Ahora daremos una breve definición de cada una de los nuevos conceptos que incorporamos al estudiar acerca de censos.



## Definiciones

### Censo:

El término censo no solo se aplica a aquellos relevamientos que comprenden todas las unidades de un país y que se realizan con una frecuencia de recolección quincenal o decenal, como es el caso de los censos de población, económicos, agropecuarios, etc., sino también a todo relevamiento, cualquiera sea su cobertura geográfica, número de unidades de información, o frecuencia de su recolección, siempre que incluya todas las unidades que componen el universo que se investiga. [11]. Usualmente se divide a la unidad del país a censar de la siguiente manera.

### Región:

Define el área sobre la que se realizó o se realizará un censo, usualmente el censo se efectúa por casco urbano.

### Fracción:

Define una subdivisión de la región, se encuentra conformada por radios censales.

### Radio Censal:

Define una subdivisión de las fracciones definidas para realizar el censo. Cada radio censal esta conformado por manzanas y son utilizados en los censos como la unidad mínima que será asignada a un censista o encuestador el cual deberá recorrer cada casa de las manzanas comprendidas en los radios censales que le fueron asignados.

### Población:

Es el total del conjunto de objetos de los cuales se requiere obtener información. Aquí el término población tiene un significado mucho más amplio que el usual, ya que puede referirse a personas, cosas, actos, áreas geográficas e incluso al tiempo.

La población debe estar perfectamente definida en el tiempo y en el espacio, de modo que ante la presencia de un potencial integrante de la misma, se pueda decidir si forma parte o no de la población bajo estudio. Por lo tanto, al definir una población, se debe cuidar que el conjunto de elementos que la integran quede perfectamente delimitado. [11]

### Censo poblacional:

Conjunto de operaciones que consisten en recopilar, resumir, valorar, analizar y publicar los datos de carácter demográfico, cultural, económico y social de todos los habitantes del país y de sus divisiones político-administrativas, referidos a un momento o período dado. Esta operación va dirigida a todas las personas que residen en viviendas, ya sean familiares



o colectivas.

La unidad básica del Censo de Población es la persona residente, pero su identificación no se hace individualmente, sino que es preciso considerar sus relaciones de convivencia, por lo que la familia, el hogar y el núcleo familiar aparecen también como unidades básicas adicionales.[10]

Censo Agropecuario:

Proporciona datos estadísticos correspondientes a las unidades de producción agropecuaria y los productores existentes en el territorio nacional. [10]

Censo Viviendas:

Conjunto de operaciones tendentes a recopilar, resumir, valorar, analizar y publicar los datos relativos a todos los lugares destinados a habitación humana que han sido concebidos como tales, y a enumerar aquellos que no habiendo sido concebidos para este fin tienen sin embargo ese uso.[10]

Residente:

Persona física que en el momento censal tiene su residencia habitual.

Vivienda:

Recinto estructuralmente separado e independiente que, por la forma en que fue construido, reconstruido, transformado o adaptado, está concebido para ser habitado por personas o, aunque no fuese así, constituye la residencia habitual de alguien en el momento censal. Como excepción, no se consideran viviendas los recintos que, a pesar de estar concebidos inicialmente para habitación humana, en el momento censal están dedicados totalmente a otros fines (por ejemplo, los que estén siendo usados exclusivamente como locales).

Un recinto se considera separado si está rodeado por paredes, muros, tapias, vallas, se encuentra cubierto por techo, y permite aislar a una persona, o un grupo de personas de otras, con el fin de preparar y consumir sus alimentos, dormir y protegerse contra las inclemencias del tiempo y del medio ambiente.

Se considera independiente si tiene acceso directo desde la calle o terreno público privado, común o particular, o bien desde cualquier escalera, pasillo, corredor, es decir, siempre que los ocupantes de la vivienda puedan entrar o salir de ella sin pasar por ningún recinto ocupado por otras personas. En todo caso, se tiene en cuenta la situación actual del recinto-vivienda y no el estado primitivo de construcción, de modo que en las agregaciones o subdivisiones de viviendas se consideran cuantas unidades hayan resultado del proceso

de transformación, siempre que cumplan las condiciones anteriormente definidas, e independientemente, por tanto, de su situación inicial de construcción.

#### Vivienda Colectiva:

Vivienda destinada a ser habitada por un grupo de personas sometidas a una autoridad o régimen común no basados en lazos familiares ni de convivencia. La vivienda colectiva puede ocupar sólo parcialmente un edificio o, más frecuentemente, la totalidad del mismo. Ej. conventos, cuarteles, asilos, residencias de estudiantes o de trabajadores, hospitales, prisiones, hoteles, pensiones y establecimientos análogos.

Cuando dentro del establecimiento colectivo existan viviendas de carácter familiar, éstas serán censadas, como viviendas familiares.

#### Vivienda Familiar:

Vivienda destinada a ser habitada por una o varias personas, general pero no necesariamente unidas por parentesco, y que no constituyen un colectivo, según la definición anterior.

#### Alojamiento:

Vivienda familiar que presenta la particularidad de ser móvil, semipermanente o improvisada, o bien que no ha sido concebida en un principio con fines residenciales pero, sin embargo, constituye la residencia de una o varias personas en el momento del censo.

Los alojamientos pueden ser fijos, como:

Las viviendas semipermanentes que, aunque son semejantes a las viviendas familiares en algunos aspectos, son utilizadas nada más que durante un tiempo limitado (normalmente menos de 10 años).

Determinados recintos destinados a vivienda que se construyen sin ningún pilar y con materiales de desecho (latas, cajas...).

Otros recintos, cuyo destino es distinto del de vivienda, y que no han sido reconstruidos ni reformados para ser utilizados con fines residenciales, pero en los que a pesar de eso vive gente en su interior; por ejemplo, los espacios situados en cuadras, pajares, molinos, garajes, almacenes, locales, así como las cuevas y otros refugios naturales que han sido habilitados para vivir en ellos.

#### Hogar:

Grupo de personas residentes en la misma vivienda familiar.

Familia:

Grupo de personas que, residiendo en la misma vivienda familiar, están vinculadas por lazos de parentesco, ya sean de sangre o político, e independientemente de su grado.

Las diferencias entre hogar y familia son:

- a) El hogar puede ser unipersonal, mientras que la familia tiene que constar, por lo menos, de dos miembros.
- b) Los miembros de un hogar multipersonal no tienen necesariamente que estar emparentados, mientras que los miembros de una familia sí.

Núcleo Familiar:

Corresponde a una concepción restringida de la familia, limitada a los vínculos de parentesco más estrechos. Existen cuatro tipos de núcleo familiar:

- a) Matrimonio o pareja sin hijos.
- b) Matrimonio o pareja con uno o más hijos.
- c) Padre con uno o más hijos.
- d) Madre con uno o más hijos.

Para formar parte del núcleo, un hijo debe ser soltero y estar no emparejado.

Local:

Recinto estructuralmente separado e independiente que no está exclusivamente dedicado a vivienda familiar y en el que se llevan o se pueden llevar a cabo actividades económicas dependientes de una empresa o institución. El recinto debe estar situado en un edificio, ocupándolo total o parcialmente.

Se entiende por actividad económica toda acción productora que llevan a la creación de bienes o la prestación de servicios. Las actividades pueden realizarse con o sin fines de lucro.

Edificio:

Construcción permanente, separada e independiente, concebida para desarrollar cualquier actividad.

Una construcción es permanente si ha sido concebida y construida para atender necesidades de duración indefinida y que, por lo tanto, durará normalmente en el mismo sitio más de diez años. Es separada si está limitada por fachadas o medianeras y cubierta por techo. Es independiente si tiene acceso directo desde la calle o desde terreno público o privado. El acceso a un edificio puede realizarse por una entrada principal e independiente, o por otras entradas secundarias o accesorias.

En el caso de construcciones que son conjuntos de bloques o edificios adosados, o cercados por un cerramiento común, se debe considerar que existen tantos edificios como entradas o portales principales e independientes posean.

La independencia del acceso se refiere a la imposibilidad de acceder a otras viviendas distintas de aquellas a las que corresponde la entrada principal. Se exceptúan las posibles comunicaciones que pueden existir por el garaje, terrazas o trasteros comunes, así como aquellas otras comunicaciones excepcionales que puede haber en algunas plantas para casos de incendio, avería de ascenso.

Si una construcción tiene una sola entrada y hay varias escaleras distintas que dan acceso a otros tantos grupos de viviendas no comunicadas entre sí salvo a través de la planta baja, se considera como un solo edificio. Las construcciones que tienen entradas a distintos niveles por dos calles diferentes, estando sus viviendas comunicadas entre sí a partir de un nivel determinado, forman un solo edificio.

#### Complejo de edificios:

Conjunto de edificios ubicados en un área limitada (cercada o no) y que se utilizan, exclusiva o principalmente, bien para la realización de las distintas fases, operaciones o necesidades de la actividad económica de un único organismo, entidad o empresa, o bien para vivienda colectiva.

No se considera complejo de edificios a un conjunto de éstos destinados exclusiva o principalmente a vivienda familiar.

A efectos censales, cada complejo de edificios se contabilizará como una unidad.

#### Unidad de análisis:

Es el objeto del cual se desea obtener información. Muchas veces nos referimos a las unidades de análisis con el nombre de elementos. En estadística, un elemento o unidad de análisis puede ser algo con existencia real, como un automóvil o una casa o algo más abstracto como la temperatura o un intervalo de tiempo. Dada esta definición, puede redefinirse población como el conjunto de unidades de análisis.

#### Muestra:

Es un subconjunto de unidades de análisis de una población dada destinado a suministrar información sobre la población. Para que este subconjunto de unidades de análisis sea de utilidad estadística, deben reunirse ciertos requisitos en la selección de los elementos.

Las causas por las que se seleccionan muestras son muchas. Puede ocurrir que la población que se defina tenga tamaño infinito, y en consecuencia, no fuera posible observar

a todos sus elementos. En otras ocasiones, el costo de la observación exhaustiva puede ser muy elevado, el tiempo de recolección de la información muy extenso, o más aun, la observación de los elementos puede ser destructiva.

#### Variable:

Es la cualidad o cantidad medible que se estudia de las unidades de análisis y que varían de una unidad a otra. Por ejemplo: la edad, ingreso de un individuo, sexo, etc.

### **La Demografía**

La demografía es una de las ciencias sociales que estudia los acontecimientos que ocurren a los miembros de una población a lo largo de su vida.

Tiene dos dimensiones:

La medición ¿Cuántos hay?, ¿Cuántos nacen?, ¿Cuántos trabajan?, ¿Cuál es su ingreso?, ¿Quiénes mueren?.

La explicación ¿por qué se tienen más hijos en La Plata que en Buenos Aires? ¿Por qué recibimos emigrantes? ¿Por qué hoy mucha más gente llega a vivir 80 años?.

#### *Análisis Demográfico*

A menudo en la búsqueda de explicaciones la demografía utiliza la economía, la historia, la biología. Por otro lado las herramientas de medición de los demógrafos y sus análisis son también aplicados al resto de ciencias que estudian al hombre, por ejemplo la estructura por edades afecta al consumo y al resto de magnitudes económicas, a la política, a las costumbres.

En general el trabajo de los demógrafos se limita dentro de la definición anterior en dos sentidos: se presta especial importancia a los agregados, a lo macro frente a lo micro; existe un núcleo central de temas que hacen referencia a la dinámica de la población:

La fecundidad interesa más que el desempleo porque la primera determina directamente cómo cambia la población. Eso no quiere decir que no se pueda hacer un estudio demográfico del desempleo, ni que el desempleo no pueda ser uno de los factores explicativos de la fecundidad. A este núcleo central dentro de la demografía de medición de la dinámica de las poblaciones se le suele denominar Análisis Demográfico. [9]

### **Tamaño de la Población**

Cómo cambia el tamaño de la población objeto de estudio es, por tanto, tema central de la demografía.

El estudio de la dinámica de una población tiene mucho de contabilidad: si partimos de una población inicial para saber cual es la población final tenemos que seguir las entradas y las salidas de la población.

A una población se puede entrar de dos formas:

- 1.- Por el nacimiento
- 2.- Por la inmigración.

De una población se puede salir de dos formas:

- 1.-Por la muerte
- 2.-Por la emigración.

De aquí se tiene la igualdad demográfica básica o Ecuación Fundamental de la Población que podemos escribir de diversas maneras:

$$\text{Población Final} = \text{Población Inicial} + \text{Nacimientos} - \text{Defunciones} + \text{Inmigración} - \text{Emigración}$$

$$\text{Población Final} = \text{Población Inicial} + \text{Crecimiento Natural ó Vegetativo} + \text{Migración Neta}$$

$$\text{Crecimiento de la Población} = \text{Nacimientos} - \text{Defunciones} + \text{Migración Neta}$$

$$P(t+1) - P(t) = DP(t) = N(t) - D(t) + I(t) - E(t)$$

Donde:

$P(t)$  es la población a principio del año  $t$ .

$N(t)$  a los nacimientos registrados en ese año.

$D(t)$  a las defunciones.

$E(t)$  e  $I(t)$  a las migraciones tenemos que en un año.

En estas igualdades se basan los índices e indicadores que producen las bases de datos primarias o secundarias, en el siguiente punto se analiza como se derivan resultados importantes de un censo a partir de la Ecuación Fundamental de la Población.

## Indicadores Demográficos

Naturalmente cuantos más habitantes tiene una población mayor será el número de nacimientos, defunciones y demás. Para hacernos una idea de lo grande que es la natalidad, la mortalidad o las migraciones tenemos que poner en relación los acontecimientos vitales con el total de población. [9]

Se definen así los Indicadores Demográficos que es un parámetro que caracteriza la información obtenida en un censo. Este trabajo se focaliza en aquellos indicadores que basan sus cálculos en información poblacional acerca de natalidad, fecundidad, mortalidad, esperanza de vida, ingresos y otros datos relacionados con la misma. A continuación damos una clasificación de los Indicadores demográficos que son de interés para esta investigación.

Las tasas brutas de natalidad, mortalidad, inmigración y emigración dividiendo el número de sucesos por la población media del período.

$$PM(t) = \frac{1}{2}[P(t+1) + P(t)].$$

Así tenemos, por ejemplo:

Tasa bruta de Natalidad:

$$TBN(t) = b(t) = N(t)/PM(t)$$

Tasa bruta de Mortalidad:

$$TBM(t) = d(t) = D(t) / PM(t)$$

Tasa bruta de inmigración:

$$TBI(t) = I(t) / PM(t)$$

Todos estos indicadores son proporciones. En la práctica suelen multiplicarse por mil para expresarse así en tantos por mil.

La tasa de crecimiento natural ó vegetativo se define por la diferencia entre la tasa bruta de natalidad y de mortalidad:

$$r(t) = b(t) - d(t)$$

Esta tasa suele habitualmente expresarse en tantos por ciento en vez de en tantos por mil (es decir, suele multiplicarse por cien).

La tasa bruta de migración neta definida como la diferencia entre las tasas de inmigración y emigración:

$$TBMN(t) = TBI(t) - TBE(t)$$

Volviendo a la ecuación de crecimiento nos damos cuenta de que podemos ahora caracterizar el crecimiento de la población sumando la tasa de crecimiento natural y la tasa bruta de migración neta. De este modo se obtiene la tasa de crecimiento de la población.

De modo similar podemos definir la tasa bruta de nupcialidad y otros indicadores basándonos en los resultados obtenidos del censo.



Estos resultados los obtienen las agencias dedicadas a realizar los censos de ellas depende que esta información al ser publicada no viole el secreto estadístico.

## Estructura por Edades

En el punto anterior se ha tomado en cuenta la población en su conjunto. A menudo es importante clasificar la población en subgrupos más homogéneos. La clasificación por edades y sexos es la más utilizada puesto que gran parte del comportamiento demográfico está condicionada por estas dos dimensiones. La representación gráfica en la que clasificamos la población de acuerdo a la edad y el sexo recibe el nombre de pirámide de población, pues para muchas poblaciones (aquellas que crecen o que tienen alta mortalidad) el aspecto del gráfico se asemeja al de una pirámide: hay menos personas según aumenta la edad.

La forma de la pirámide de la población es importante, pues nos indica lo joven o anciana que es esta población. Si la base es muy ancha, esto quiere decir que hay muchos jóvenes. Esto tendrá implicaciones importantes para la educación o la estructura familiar, por ejemplo. Si la cúspide es muy ancha, tenemos una población madura o vieja. Esto tiene a su vez efectos sobre el gasto en salud o pensiones por ejemplo.

La pirámide de población también nos informa sobre el pasado. Las personas de una edad determinada son personas que nacieron en un mismo período, es decir: cada franja de la pirámide de población representa a una generación o cohorte. El número de nacidos inicialmente representaría el tamaño de la cohorte. Según pasan los años van falleciendo miembros de la cohorte de modo que si seguimos a una generación en una secuencia de pirámides de población, vamos viendo como su tamaño va disminuyendo. Esto ocurriría en ausencia de migraciones. Como las migraciones suelen afectar más a los jóvenes, la existencia de fuertes migraciones modifica la forma de la pirámide de población. Esto es particularmente evidente cuando la población que analizamos es relativamente reducida, por ejemplo, una ciudad o un pueblo.

## Análisis de Cohorte

Hemos visto que una cohorte de nacimiento, entendiendo como tal un conjunto de personas nacidas en un período determinado, puede ser seguida a lo largo de su vida. Si seguimos a todas estas personas a lo largo de su vida hasta que todos han muerto puedo resumir cuál ha sido el comportamiento demográfico de la cohorte. A este tipo de análisis se le llama análisis longitudinal o de cohorte.

Para resumir como de intensa ha sido la mortalidad de la cohorte estudiamos lo longevos que han sido sus miembros. En particular, es posible calcular cuántos años han vivido de media. Para ello se suman las edades exactas al fallecimiento de todos ellos y se divide



por el tamaño de la cohorte. A esta cantidad se le denomina esperanza de vida de la cohorte.

Para resumir como se ha reproducido la cohorte ponemos en relación el número de hijos que han tenido con el tamaño de la cohorte. Es práctica habitual centrarse únicamente en la población de mujeres.

El número de hijas dividido entre el número de mujeres inicialmente en la cohorte es la razón de reproducción neta o tasa de reproducción neta (*NRR*, Net Reproduction Ratio en inglés). Un número mayor que uno indica que la generación ha contribuido al aumento de la población.

La razón de reproducción neta no es estrictamente una medida de la fecundidad. Para tener hijas es necesario que las mujeres hayan sobrevivido hasta el momento del nacimiento, y no todas las mujeres de la cohorte lo consiguen. La mortalidad, por lo tanto, también influye sobre la *NRR*. Es posible, sin embargo, eliminar el efecto de la mortalidad y estimar cuántos hijos de media habría terminado teniendo cada mujer de la cohorte en ausencia de mortalidad. Para ello calculamos las tasas de fecundidad de la cohorte a cada edad, poniendo en relación el número de hijos con las mujeres supervivientes. A la suma de estas tasas se le denomina descendencia final.

La utilidad del análisis longitudinal no se restringe al puro análisis demográfico. En particular el análisis longitudinal es muy aplicable a la economía: podemos observar a qué edad los distintos miembros de una cohorte dejan de estudiar, comienzan a trabajar, los ingresos que tienen, se compran una vivienda o cualquier otra dimensión que nos interese.

## Dinámica Demográfica

El análisis longitudinal o de cohortes es muy útil para describir generaciones. No lo es tanto para describir la evolución de la población. Como hemos visto una población en un instante está compuesta por distintas generaciones observadas cada una de ellas en una edad diferente. El número de nacimientos, el número de defunciones será finalmente el resultado de una combinación de circunstancias: el comportamiento y características de cada una de las cohortes por un lado, y la estructura por edades de la población por otro, que determina el peso relativo de las distintas generaciones en el total. El análisis de la población en un momento dado recibe el nombre de análisis transversal o de período o del momento.

Si queremos predecir como va a cambiar una población en un momento dado, la información que nos da la estructura por edades es importante: no todo el mundo tiene el mismo riesgo de morir. Es mucho más probable morir para un anciano que para un joven.

Tampoco todas las personas pueden tener hijos. Sólo las mujeres en edad fértil pueden dar a luz un hijo. Por este motivo para comprender la dinámica demográfica no basta con conocer las tasas brutas:

Una tasa bruta de mortalidad baja puede ser debida a que los riesgos de morir sean muy bajos o a que la población sea muy joven. Por este motivo construimos las tasas específicas de mortalidad por edades. Estas se construyen dividiendo las personas muertas a una determinada edad en un año por la población media de esa edad:

$$mx(t) = Dx(t) / PMx(t)$$

donde  $x$  indica la edad.

En la práctica estas tasas se calculan de forma separada para hombres y mujeres puesto que las tasas son distintas para los dos sexos.

A partir de las tasas específicas de mortalidad por edades es posible construir una tabla de mortalidad de período. Esta consiste en un ejercicio de imaginación que se conoce como análisis de cohorte sintética o ficticia:

Si se toma una cohorte supuesta de tamaño arbitrario, por ejemplo, 10000 individuos, y calculamos cuántos morirían a cada edad de esa cohorte si sus tasas de mortalidad son las tasas específicas del momento. Al número de muertes a cada edad de esa cohorte ficticia le denominamos  $dx$ . Al número de supervivientes de esos 10000 mil que llegarían a cumplir  $x$  años se le llama  $lx$ . Ambas son funciones de  $x$  que nos sirven para caracterizar la mortalidad en un instante dado. Es también posible calcular la esperanza de vida al nacimiento de la cohorte ficticia  $e0(t)$ : cuántos años vivirían de media unos individuos que experimentarían a lo largo de su vida las tasas de mortalidad por edades del momento actual. Esta es la medida pura de mortalidad de uso más común. Es importante señalar que es una medida de período: no responde a la experiencia de una cohorte real. Es errónea su interpretación, muy común por otra parte, como el número esperado de años que vivirá un recién nacido de ese año. Esta última cantidad depende de la mortalidad futura, y se corresponde con la esperanza de vida de la cohorte de nacidos en ese año. Sin embargo, esta sólo podrá ser conocida con certeza cuando todos los miembros de esa cohorte hayan fallecido, a posteriori.

El artificio de la cohorte sintética se utiliza también para calcular indicadores de fecundidad puros. Para ello calculamos las tasas específicas de fecundidad por edades.

$$fx(t) = Nx(t) / PMx(t)$$

donde:

$Nx(t)$  es el número de nacimientos en el período cuyas madres tenían edad  $x$ .

$PMx(t)$  es la población media de mujeres de edad  $x$ .

Puede interpretarse como la proporción de mujeres de esa edad que tienen hijos en ese año. Sumando las tasas específicas de fecundidad por edades obtenemos el equivalente en análisis de período a la descendencia media: el Índice Sintético de Fecundidad o Tasa de Fecundidad Total (*ISF* ó *TFR* de sus iniciales en inglés, Total Fertility Rate). Se trata del indicador de fecundidad de período más utilizado, y se expresa en hijos por mujer. Representa el número medio de hijos que acabaría teniendo una mujer de la cohorte sintética (de ahí lo de índice sintético) que sobreviviera las edades fértiles y tuviera hijos de acuerdo a las tasas específicas de fecundidad por edades del momento actual.

Finalmente es también posible combinar fecundidad y mortalidad y estudiar la reproducción en la cohorte sintética, es decir: la razón de reproducción neta de período. Para ello no tendríamos más que aplicar simultáneamente las tasas de mortalidad y de fecundidad a la cohorte sintética de mujeres y dividir el número de hijas que tuvieran por el número inicial de mujeres. Esta medida me puede dar una idea de las implicaciones a largo plazo de las tasas de mortalidad y fecundidad del momento. Un valor de 2 indicaría que si se mantienen las tasas actuales la población tendería a duplicarse cada generación. Un valor de 1 indicaría que la población tendería a perpetuarse en un tamaño concreto. Por eso decimos que 1 es el nivel de reemplazo de la Razón de Reproducción Neta. A menudo también se habla del nivel de reemplazo del Índice Sintético de Fecundidad, es decir, cuántos hijos tendrían que tenerse por mujer en la cohorte sintética para que la *NRR* fuera igual a 1. Esta cantidad depende del nivel de mortalidad, pero suele estar en torno a 2,1 que suele indicarse como la fecundidad de reemplazo. Como vemos el análisis demográfico de período lleva de forma natural a preguntarse sobre las implicaciones a largo plazo de unas tasas concretas de mortalidad y fecundidad. Esto se corresponde con la idea de las poblaciones estables.

## Poblaciones Estables

Es uno de los núcleos centrales del análisis demográfico, es el estudio de qué ocurre cuando se mantienen fijas a lo largo del tiempo la fecundidad, representada por las tasas de fecundidad por edades  $fx$ , y la mortalidad, representada por las tasas de mortalidad por edades  $mx$ , o más comúnmente por la función de supervivientes de la tabla de mortalidad  $lx$  (es posible pasar de una a otra). A una población que cumple estas condiciones a lo largo del tiempo se le denomina Población Estable. [7]

La respuesta a qué ocurrirá si las tasas no cambian es que, a largo plazo, la población tenderá a crecer a una tasa de crecimiento constante, la tasa de crecimiento intrínseca de esa población  $r$ . Además, la estructura por edades, independientemente de cuál fuera la estructura por edades inicial de la población, tenderá a acercarse a una estructura fija. Es

la llamada estructura por edades de la población estable. Una vez que una población alcanza su estructura estable, ésta no cambiará a lo largo del tiempo, es decir: la población podrá crecer al ritmo indicado por  $r$ , pero la proporción de personas de una edad concreta en la población no cambiará.

La tasa de crecimiento intrínseca está relacionada con la razón de reproducción neta. Una primera aproximación de esta se obtiene como

$$r = \log(NRR) / m$$

donde:

$m$  es la edad media a la maternidad.

Por otro lado es también común relacionar  $r$  con el Índice Sintético de Fecundidad.

$$ISF = lm \times NRR$$

donde:

$lm$  representa la proporción de mujeres que sobreviven hasta la edad media al tener hijos.

A su vez esta proporción guarda en la práctica una relación aproximada con la esperanza de vida al nacer  $e_0$ . En definitiva, se puede escribir  $r$  como una función aproximada del Índice Sintético de Fecundidad  $ISF$ , la esperanza de vida  $e_0$ , y la edad media a la maternidad  $m$ .

### Ejemplos

Los indicadores que se muestran a continuación provienen de datos censales, estadísticas del Ministerio de Salud, estimaciones y proyecciones demográficas para los períodos 1950-2050 y 1990-2010. Dichas estimaciones fueron elaboradas a partir de los resultados del Censo de Población 1991[8].

Los valores de 1869 y 1895 corresponden a los dos primeros períodos Intercensales, mientras que los valores sucesivos provienen de tablas de vida elaboradas para cada fecha censal.

Indicadores demográficos por provincia. Años 1990-2001[8].

Provincia	Tasa anual media de Crecimiento por mil 1991/2001(1)	Esperanza de vida al nacer 1990-92 en años			Tasa bruta de natalidad por mil 1999	Tasa bruta de Mortalidad por mil 1999	Tasa de Mortalidad infantil por mil nacidos vivos 1999
		Total	Varon	Mujer			
Total del país	11,0	71,93	68,44	75,59	18,8	7,9	17,6
Ciudad de Buenos Aires	-6,8	72,72	69,17	76,45	13,5	11,8	10,7
Buenos Aires	9,3	72,09	68,53	75,78	17,4	8,3	16,6
19 partidos del Gran Buenos Aires	10,8	.	.	.	.	.	.
Resto Buenos Aires	6,4	.	.	.	.	.	.
Catamarca	23,6	70,61	67,40	73,96	24,9	6,1	20,3
Chaco	15,9	69,02	65,64	72,55	24,1	6,6	29,0
Chubut	14,7	70,58	67,26	74,04	18,9	5,4	17,9
Córdoba	10,2	72,79	69,15	76,60	17,5	8,4	15,3
Corrientes	15,6	70,09	67,16	73,16	23,4	6,2	22,9
Entre Ríos	12,6	71,61	68,13	75,26	20,1	8,4	19,5
Formosa	20,0	69,37	66,27	72,62	25,3	5,5	23,6
Jujuy	17,8	68,37	65,24	71,65	23,0	5,6	23,4
La Pampa	13,9	71,57	68,15	75,15	18,5	7,3	15,2
La Rioja	27,6	70,38	67,04	73,89	22,3	5,8	20,6
Mendoza	11,1	72,72	69,80	75,75	19,0	7,0	16,2
Misiones	20,2	69,49	66,45	72,65	24,5	5,1	21,7
Neuquén	19,9	71,39	68,30	74,67	18,5	3,7	12,4
Río Negro	8,7	70,87	67,54	74,36	18,4	4,9	15,8
Salta	22,3	68,92	66,13	71,84	24,8	5,3	21,2
San Juan	16,4	71,13	68,10	74,30	22,3	7,2	20,7
San Luis	25,1	70,79	67,64	74,07	22,2	6,5	18,9
Santa Cruz	21,1	70,41	67,11	73,90	20,3	4,5	13,2
Santa Fe	6,9	72,29	68,50	76,28	17,3	8,7	15,2
Santiago del Estero	18,4	69,83	67,13	72,73	21,1	6,4	15,5
Tierra del Fuego, Antártida e Islas del Atlántico Sur							
(2)	38,2	70,16	66,85	73,64	18,7	2,5	7,8
Tucumán	15,9	71,01	68,08	74,13	21,6	6,8	22,5

Tabla 3.1 Tasa anual media de crecimiento, esperanza de vida, natalidad, mortalidad y mortalidad infantil.

(1) Es el número medio de personas que se incorporan anualmente a la población total, por 1.000 habitantes, por cuanto la tasa de crecimiento geométrico es:



Siendo:

$$P(t) = \text{Población al final del período}$$

(según Censo 2001);

$$r = \sqrt{\frac{P(t) - P(0)}{t}}$$

$$P(0) = \text{Población al inicio del período}$$

(según Censo '91);

$t = \text{tiempo transcurrido entre ambos censos.}$

(2) No incluye Sector Antártico e Islas del Atlántico Sur.

Estimación de indicadores demográficos. Años 1950-2050 la elaboración de estos indicadores está basada en las proyecciones de población realizadas a partir de los resultados definitivos del Censo Nacional de Población 1991.[8]

Período	Tasa de reproducción (por mujer)		Tasa de fecundidad (por mil mujeres)		Edad media de la fecundidad en años	Esperanza de vida al nacer en años			Tasa de mortalidad infantil 0/00 nacidos vivos
	Bruta	Neta	Global	General		Total	Varon	Mujer	
1950-1955	1,5	1,4	3.154,0	97,6	28,3	62,7	60,4	65,1	65,9
1955-1960	1,5	1,4	3.126,5	95,4	28,3	64,7	62,1	67,4	60,4
1960-1965	1,5	1,4	3.089,5	92,7	28,3	65,5	62,5	68,6	59,7
1965-1970	1,5	1,4	3.049,0	90,6	28,3	66,0	62,8	69,3	57,4
1970-1975	1,5	1,4	3.145,5	94,5	28,0	67,4	64,1	70,8	48,1
1975-1980	1,7	1,6	3.440,0	105,5	27,7	68,8	65,4	72,2	39,1
1980-1985	1,5	1,5	3.150,0	97,0	27,8	70,2	66,8	73,7	32,2
1985-1990	1,5	1,4	3.003,0	91,8	27,7	71,0	67,6	74,6	27,1
1990-1995	1,4	1,3	2.828,0	85,7	27,6	72,1	68,6	75,7	24,3
1995-2000	1,3	1,2	2.619,0	80,1	27,5	73,1	69,7	76,8	21,8
2000-2005	1,2	1,2	2.444,0	76,2	27,4	74,1	70,6	77,7	20,0
2005-2010	1,1	1,1	2.303,0	71,9	27,4	75,1	71,6	78,7	17,9
2010-2015	1,1	1,0	2.189,0	67,2	27,3	75,9	72,4	79,5	16,0
2015-2020	1,0	1,0	2.100,0	62,8	27,3	76,7	73,3	80,4	14,3
2020-2025	1,0	1,0	2.100,0	61,5	27,3	77,5	74,0	81,1	12,8
2025-2030	1,0	1,0	2.100,0	61,1	27,3	78,2	74,7	81,8	11,4
2030-2035	1,0	1,0	2.100,0	61,1	27,3	78,8	75,3	82,4	10,2
2035-2040	1,0	1,0	2.100,0	60,6	27,3	79,4	75,9	83,0	9,1
2040-2045	1,0	1,0	2.100,0	60,1	27,3	79,9	76,4	83,5	8,2
2045-2050	1,0	1,0	2.100,0	60,0	27,3	80,4	76,9	84,0	7,3

Tabla 3.2 Tasas de reproducción, fecundidad, mortalidad infantil; edad media de la fecundidad, esperanza de vida.

Crecimiento de la población / Datos estimados desde 1970 en adelante. Proyecciones de población por sexo y grupos de edad: urbana - rural y económicamente activa (1990 - 2025) y por provincia (1990 - 2010) (Versión revisada - febrero 1996). Serie Análisis Demográfico 7.[8]

División político- territorial		Total del país	Urbana	Rural
Población (en miles)	1970	23,36	18,45	4,91
	1980	27,95	23,19	4,76
	1990	32,53	28,26	4,27
	2000	37,03	33,17	3,87
	2010	41,47	37,89	3,58
Tasa de crecimiento anual medio (‰)	1970/80	18	23	-3,2
	1980/90	15,8	20,6	-11
	1990/00	13,1	16,1	-9,9
	2000/10	11,4	13,4	-7,6
Variación relativa (%)	1970/80	19,6	25,7	-3,2
	1980/90	16,4	21,8	-10,2
	1990/00	13,8	17,4	-9,5
	2000/10	12	14,2	-7,3

Tabla 3.3 Tamaño y ritmo de crecimiento de la población urbana-rural. Total del país. 1970-2010.

Indicadores de dinámica demográfica y sus componentes. Total país. Años 1970-2010.[8]

Indicador (1)	1970-1975	1975-1980	1980-1985	1985-1990	1990-1995	1995-2000	2000-2005	2005-2010
Tasa de crecimiento total (%)	16,8	15,2	15,3	14,3	13,4	12,7	12,0	10,8
Tasa de crecimiento natural (‰)	14,4	16,7	14,6	13,4	12,6	11,9	11,3	10,4
Fecundidad								
Tasa bruta de natalidad (‰)	23,4	25,7	23,1	21,8	20,8	19,9	19,1	18,0
Tasa global de fecundidad	3,2	3,4	3,2	3,0	2,8	2,6	2,4	2,3
Tasa neta de reproducción	1,4	1,6	1,5	1,4	1,3	1,2	1,2	1,1
Mortalidad								
Tasa bruta de mortalidad (%)	9,0	8,9	8,5	8,5	8,2	8,0	7,8	7,7
Esperanza de vida al nacer								
Ambos sexos	67,37	68,76	70,21	71,03	72,08	73,13	74,12	75,05
Varones	64,10	65,44	66,82	67,58	68,60	69,65	70,64	71,57
Mujeres	70,78	72,22	73,74	74,62	75,70	76,75	77,74	78,67
Años de sobrevivencia femenina	6,68	6,78	6,92	7,04	7,10	7,10	7,10	7,10

Tabla 3.4 Tasas de: crecimiento total, natalidad, fecundidad, mortalidad; y datos sobre esperanza de vida

Tasa global de fecundidad según división político-territorial. Años 1990-2010

División Político- Territorial	Tasa global de fecundidad			
	1990-1995	1995-2000	2000-2005	2005-2010
<b>Total del país</b>	<b>2,83</b>	<b>2,62</b>	<b>2,44</b>	<b>2,30</b>
Capital Federal	1,70	1,58	1,47	1,39
Buenos Aires	2,67	2,47	2,30	2,17
Catamarca	3,46	3,20	2,99	2,82
Córdoba	2,66	2,46	2,30	2,16
Corrientes	3,47	3,22	3,00	2,83
Chaco	3,64	3,37	3,15	2,97
Chubut	3,16	2,93	2,73	2,57
Entre Ríos	3,09	2,86	2,67	2,52
Formosa	3,82	3,54	3,30	3,11
Jujuy	3,39	3,14	2,93	2,76
La Pampa	3,08	2,85	2,66	2,51
La Rioja	3,39	3,14	2,93	2,76
Mendoza	3,00	2,78	2,59	2,44
Misiones	3,86	3,58	3,34	3,15
Neuquén	3,30	3,06	2,85	2,69
Río Negro	3,19	2,96	2,76	2,60
Salta	3,52	3,26	3,04	2,87
San Juan	2,97	2,75	2,56	2,42
San Luis	3,14	2,91	2,71	2,56
Santa Cruz	3,29	3,05	2,84	2,68
Santa Fe	2,83	2,62	2,45	2,31
Santiago del Estero	3,72	3,44	3,21	3,03
Tierra del Fuego	3,51	3,25	3,03	2,86
Tucumán	2,99	2,77	2,59	2,44

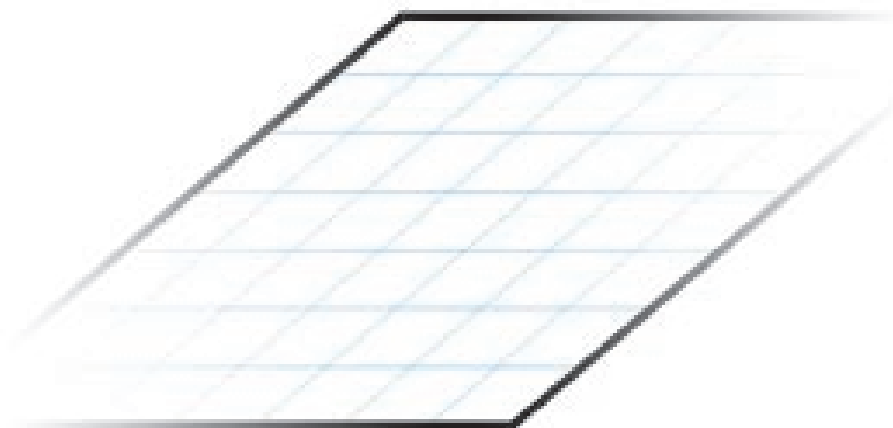
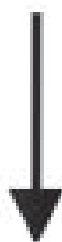
*Tabla 3.5 Tasa global de fecundidad*





# CAPÍTULO IV

## DEFINICIÓN DEL PROBLEMA



## INTRODUCCIÓN

En este capítulo explicaremos un problema real que relevamos para definir las relaciones de los objetos definidos en el Modelo resultante [24].

Nuestra intención es proporcionar un modelo sobre el cual se pueda organizar la información de tal forma que se puedan generar resultados matemáticos con facilidad, modelando la información de forma tan simple como lo que la misma representa.

Cuando hablamos de censos anteriormente, buscamos definiciones concretas de las entidades que intervienen en los mismos, en adelante utilizaremos esos conceptos para definir solamente las entidades que sean relevantes.

Explicamos la situación que plantea el problema tomando barrios de capital federal sobre mapas con diferente división geográfica y detallaremos que dificultades se le presentan a los encargados de procesar y publicar los censos.

También en este apartado fundamentaremos, el porque es necesario organizar la información sobre un modelo a la hora de controlar los resultados, en particular cuando se comparan áreas geográficas pequeñas que tienen cantidades poblacionales reducidas y su relación con preservar el Secreto Estadístico. Aunque este último punto no es relevante para el modelo si lo es para fundamentar el objetivo del mismo.

Entonces tomaremos las definiciones de los conceptos inherentes a censos, lo aprendido de lo que un SIG aporta a la hora de modelar entidades geográficas y un problema que genere preguntas que deberán responderse con facilidad.

Nuestra idea con añadir este problema a nuestro trabajo fue aclarar como se combinaran ambos temas, SIG y Censos, en el modelo que se definirá en el capítulo siguiente.

## DEFINICIÓN DEL PROBLEMA

En Argentina, como en cualquier país, existe un organismo de gobierno encargado de todo tipo de censos, este organismo además de planificar y llevar adelante un censo, debe procesar la información resultante y publicarla con las medidas de seguridad adecuadas en cuanto a privacidad se refiere, este procesamiento de la información implica tomar los datos de las encuestas y cargarlos en bases de datos sobre las que después se realizaran las consultas.

Estas consultas se muestran luego de aplicarles filtros que las distorsionen de modo tal que no divulgue nada particular, sino que sirvan para que el analista estadístico pueda sacar conclusiones en base a las mismas, conclusiones como el nivel de mortalidad, natalidad, etc.

Como mencionamos en el capítulo III estas consultas se llaman Indicadores demográficos, la dificultad que muestra este problema ocurre cuando se procesa un indicador demográfico de una región pequeña con poca población para un periodo que abarca más de un censo, recordemos que los censos se realizan cada 10 años y su planificación en cuanto a que división geográfica tomará, puede variar de un censo al otro, lo cual producirá intersecciones en el momento de su procesamiento sobre las que se deberá contar con mas grado de detalle. También puede ocurrir en el mismo censo, cuando se generan requerimientos especiales que no se pueden calcular con la división considerada en el censo y por lo tanto existe una demanda de geografías alternativas, provocando que se generen conjunto de datos para cada una de ellas, por ejemplo conjuntos adicionales para sectores postales, distritos judiciales, etc.

Los requerimientos descriptos pueden ser muchas veces complicados de resolver y no son irrelevantes, debido a que las entidades que los solicitan poseen diferentes intereses sobre el mismo grupo de población. Las entidades que los solicitan son el gobierno, los entes gubernamentales que se encargan de controlar salud, transporte, seguridad.

Estamos acostumbrados a escuchar acerca del nivel de pobreza o de la deserción escolar pero no nos cuestionamos como es que llegan a obtener los resultados que publican, o de cómo se monitorean los cambios en la población a través del tiempo.

En nuestro país el responsable de esta tarea es el INDEC, otras organizaciones también trabajan resultados que este organismo genera analizándolos con diferentes objetivos, por ejemplo los organismos que trabajan otorgando las habilitaciones para estaciones de servicio, deben poder medir factores seguridad, población, etc.

El INDEC recibe habitualmente demandas de información que no pueden ser satisfechas con los datos publicados en formato estandar. Estos trabajos especiales son aplicables a múltiples propósitos como el que ya mencionamos. Para esto se pueden solicitar nuevos procesamientos a partir de las bases de datos que la Institución posee, con el único límite que impone el Secreto estadístico. Estos trabajos especiales tienen un costo adicional por el trabajo que implica combinar en muchos casos la información para satisfacer la consulta.[8]

Al generar las estadísticas para estos trabajos especiales se debe combinar la información y esta combinación produce intersecciones debido a las diferencias entre las divisiones geográficas. Si la información no es precisa el analista estadístico no podría definir si la misma es publicable o no.

Como mencionamos anteriormente la intención es que se genere esta información fácilmente y sin error, para realizarlo existen técnicas que actualmente se aplican casi manualmente, decimos casi porque si bien es cierto que la información se carga en la base

de datos, con las herramientas actuales hay que hacer controles manuales sobre las mismas para poder definir que información es inherente para el Indicador Demográfico solicitado.

Un censo representa la información que describe la población de una región en particular. Sobre las regiones se distinguirán además las fracciones y una unidad mínima de medida, ambos son también regiones pero se les denomina así por el tamaño de región que representan. Entonces, las regiones están compuestas por fracciones y las fracciones se subdividen en una unidad mínima sobre la cual se pueden publicar datos de censos. En cuanto al tamaño de esta unidad mínima y su nombre dependen netamente del país en cuestión, en Argentina se denomina Radio Censal y debe representar al menos 3 manzanas, tuvimos la oportunidad de charlar con una encuestadora la cual nos oriento de cómo se distribuían los radios censales y que cantidad de manzanas debe incluir.[12]

Cuando mencionamos que existían filtros que se aplicaban a la información, nos referíamos justamente a aplicarlos a estas unidades mínimas, ya que son lo suficientemente pequeñas como para garantizar que los métodos de protección tengan el efecto deseado al momento de ser publicado a través de cualquier Indicador Demográfico. Tales métodos de protección incluyen entre otros intercambio de registros, redondeo, etc.

Concluyendo, la demanda de estadísticas de regiones basadas en diferentes geografías administrativas, postales, de grilla y comparables a través del tiempo representa tanto una oportunidad como una amenaza. La oportunidad esta en crear varios conjuntos de datos de regiones pequeñas que puedan ser vendidas a las entidades solicitantes y que se ajusten más precisamente a sus necesidades. La amenaza esta en generar datos de los cuales se pueda extraer información acerca de individuos identificables, para que esto no ocurra además de fijar una unidad de medida debemos fijar un mínimo de personas y de hogares para la región de la cual se están visualizando las estadísticas, sin embargo, si están publicando estadísticas de geografías distintas, sería posible restar un conjunto de datos del otro para obtener estadísticas para áreas por debajo del mínimo.

Como protección adicional se utiliza Intercambio de registros, para asegurar que un investigador no pueda tener la certeza de que las cantidades tabuladas en cero sean en realidad correctas. El peligro de diferenciar aparece cuando es posible anidar pequeñas regiones de una geografía en conjuntos de regiones de otra geografía.

## **CARACTERÍSTICAS DEL PROBLEMA**

Cuando las oficinas de censos publican datos de censos, generalmente lo hacen bajo la obligación legal de proteger la privacidad de los individuos, por lo que la preocupación aumenta en los casos en que los conjuntos de datos u otros informes se refieren a pocas personas debido a que en esos casos es más fácil distinguir individuos con combinaciones únicas de características identificatorias. Para el propósito de esta investigación la

posibilidad de aislar a un individuo con un único conjunto de características se considera un caso en el que hay riesgo de identificación.

Entendimos que existen características que podrían poner en riesgo a una persona si por ejemplo se conocen sus ingresos o posee alguna incapacidad ej. Sordomudo los cuales son presas fáciles de asaltos.

En Argentina la LEY N° 17.622 establece en el artículo 10° “... *que las informaciones que se suministren a los organismos que integran el Sistema Estadístico Nacional, en cumplimiento de la presente ley, serán estrictamente secretos y sólo se utilizarán con fines estadísticos. Los datos deberán ser suministrados y publicados, exclusivamente en compilaciones de conjunto, de modo que no pueda ser violado el secreto comercial o patrimonial, ni individualizarse las personas o entidades a quienes se refieran. Quedan exceptuados del secreto estadístico los siguientes datos de registro: nombre y apellido, o razón social, domicilio y rama de actividad...*”. Dejamos una referencia del tema para el que desee mas información sobre el mismo.[8]

La geografía es un factor importante en el Control de Divulgación Estadística, cuanto más grandes sean las regiones para las cuales se generan conjuntos de datos, mayor será la cantidad de habitantes y por lo tanto será menor el riesgo de divulgación de información de individuos específicos.

## Ejemplificamos el Problema

Ahora presentamos las dos configuraciones topológicas que venimos mencionando para una misma región en donde haya cambiado el radio censal y mostramos como puede producirse un indicador demográfico que se encuentre por debajo del límite mínimo anteriormente mencionado.

En la figura 4.1 mostramos la geografía A y en la figura 4.2 mostramos la geografía B. Si bien ambas geografías son de la misma región se encuentran subdivididas por distintos criterios. Suponemos que se cuenta con la información relevada por los censistas, para ambas geografías.

En base a esta información examinamos cada una de las regiones, aproximándonos a una fracción en particular y tomando cada radio censal, esta aproximación se muestra en las figuras 4.1.1 y 4.2.2, Luego para cada una de las geografías mostramos otro acercamiento en las figuras 4.1.2 y 4.2.2. Rotulamos al resultado de solapar las dos geografías marcando las intersecciones de acuerdo a las dos fracciones padres.

Lo que buscamos es encontrar las intersecciones, marcamos con letras las fracciones y le asignamos un número a los radios censales, vemos que A1 intersección B2, A2 intersección B1 y A2 intersección B4, A3 intersección B2, B6 intersección A5, A7

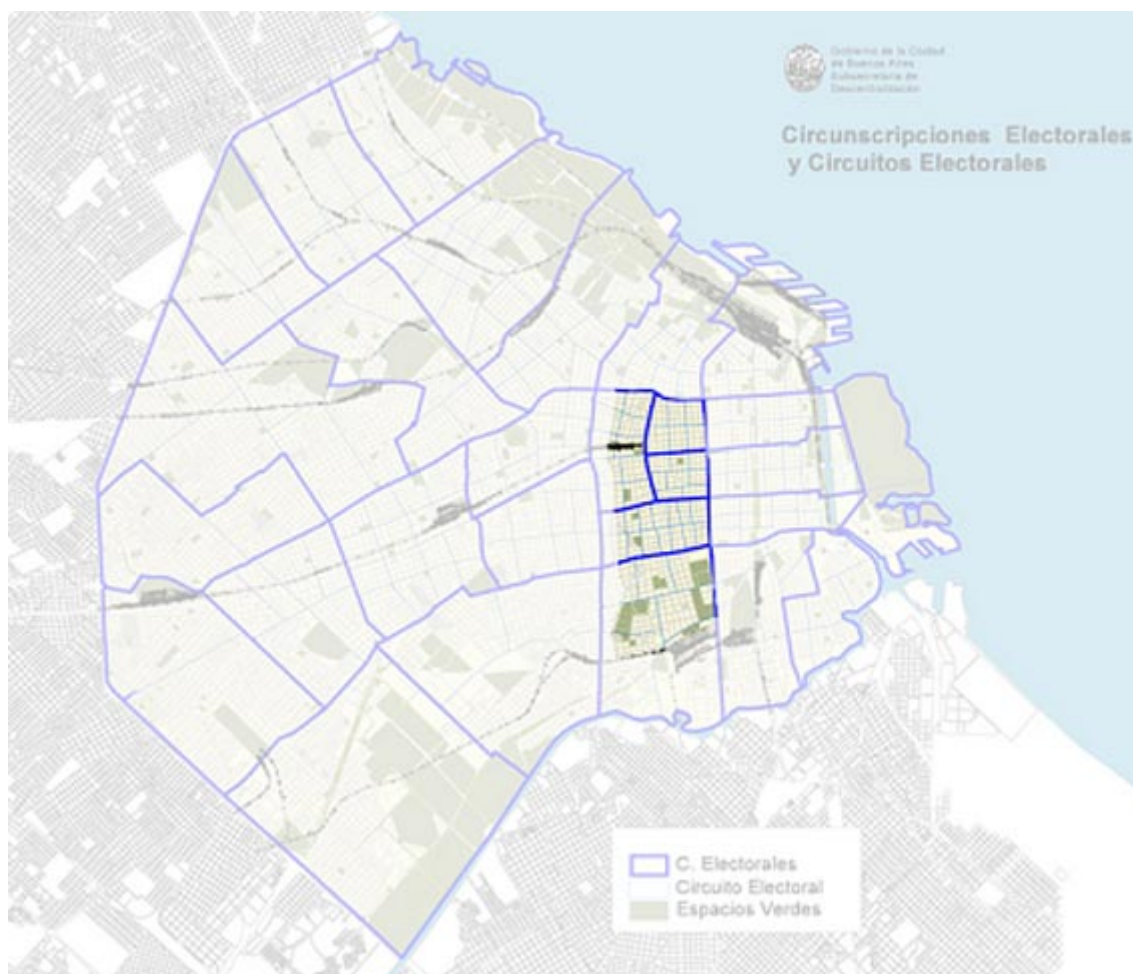
intersección B11, como mostramos en la figura 4.3.

Una vez producidos los indicadores de las geografías ingresadas los encargados de analizarlos deberán poder responder con facilidad preguntas como las siguientes ¿Cuántas de las áreas diferenciadas quedan por debajo de los límites de confidencialidad? y ¿bajo que clase de circunstancias, teniendo en cuenta las dos geografías que se comparan? entre otras.



*Figura 4.1 Capital Federal División por barrios, marcados con color verde*





*Figura 4.2 Circuitos Electorales, marcados con color azul*



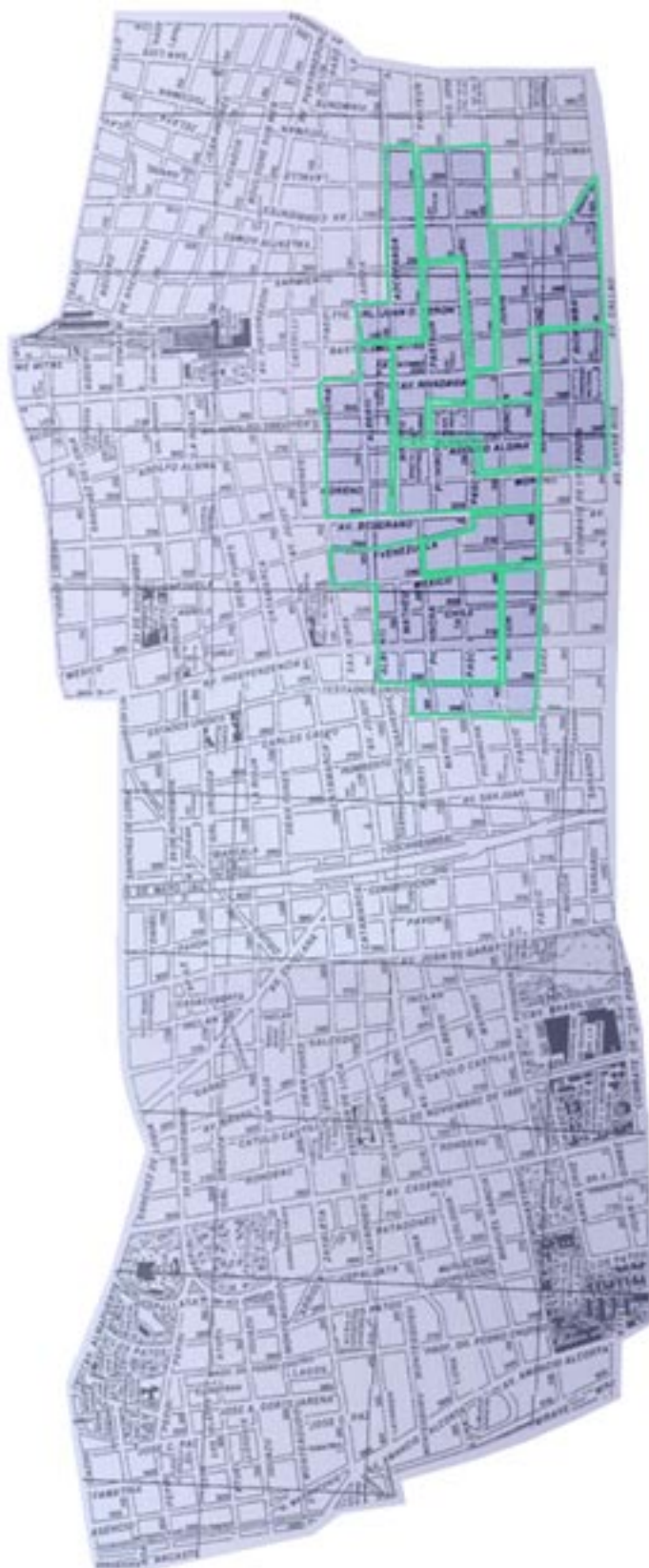


Figura 4.1.1 acercamiento de geografía A



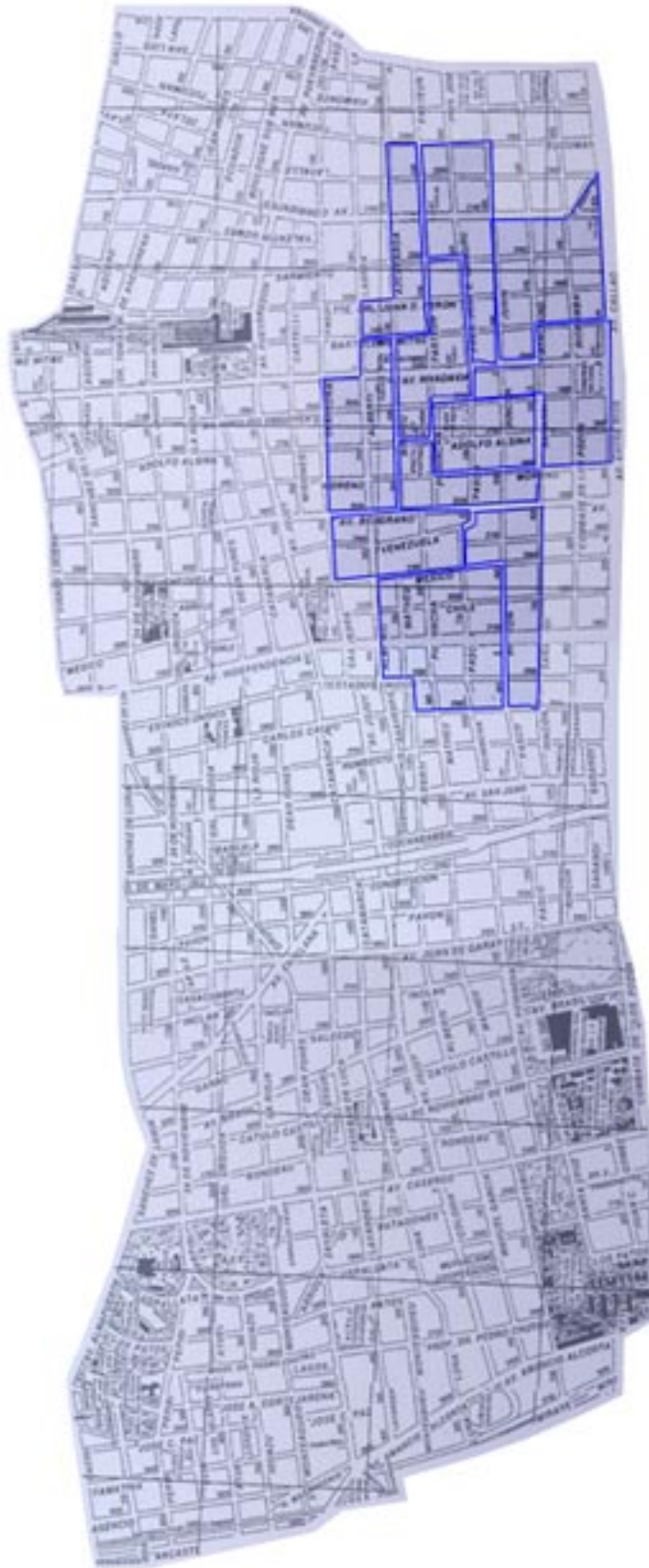


Figura 4.2.1 acercamiento de geografía B

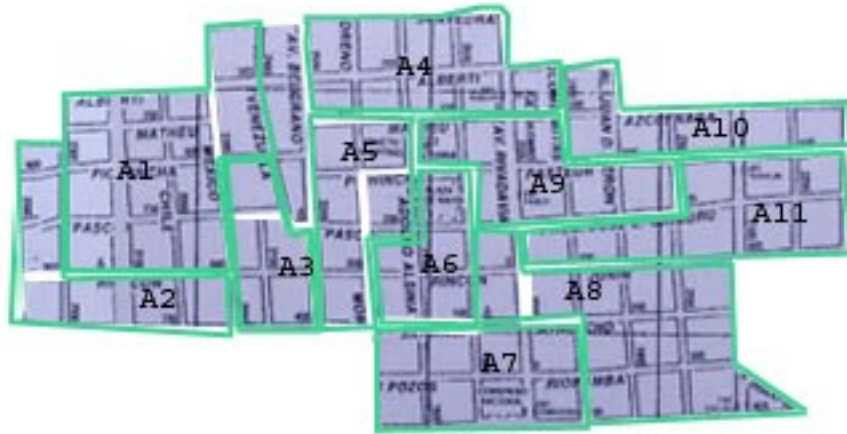


Figura 4.1.2 acercamiento de geografía A

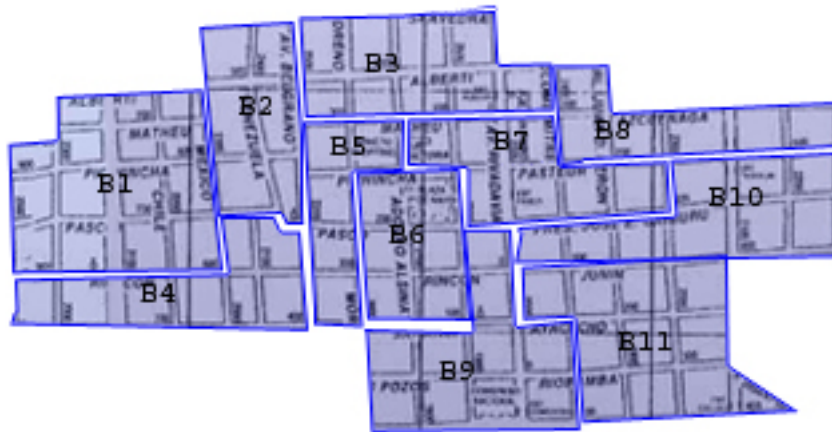


Figura 4.2.2 acercamiento de geografía B

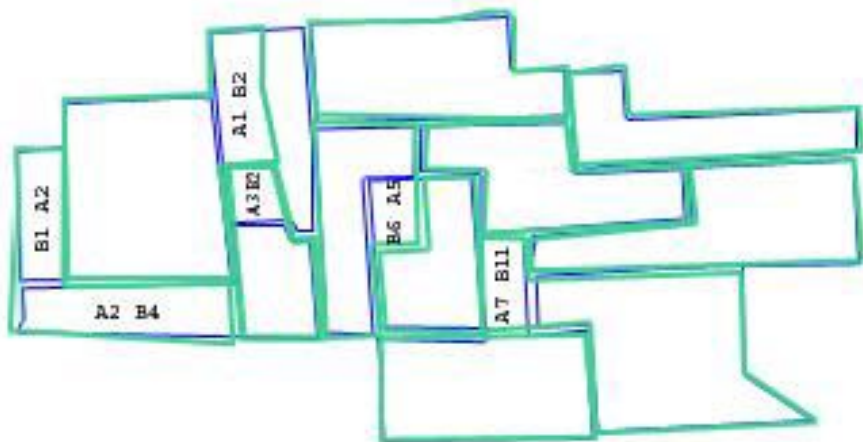


Figura 4.3 Selección de la fracción a Analizar intersecando las geografías A Y B

## CONCLUSIONES PRELIMINARES

En el ejemplo que nos fue de utilidad para determinar relaciones de las entidades que se definen en un censo, es decir lo que se espera de él. A partir del mismo empezamos a generar conclusiones y tomar características que nos servirán para diseñar nuestro modelo.

Iniciaremos esta nueva etapa de conclusiones mencionando de manera general las características que fuimos encontrando hasta el momento y que servirán de guía para iniciar el Modelo de Datos.

- Un censo poblacional se define a través de la región sobre la que se realiza y de las personas que habitan en la misma.
- Una región se encuentra definida por fracciones y las fracciones por radios censales, estos últimos están compuestos por manzanas, donde la cantidad está limitada por el usuario. Una manzana está dividida en lotes, los cuales pueden contener viviendas.
- Sobre las regiones es necesario definir su ubicación y sus características topológicas.
- Se deben poder crear los censos, hogares y personas dentro de cada región de cada geografía ingresada con un número que lo identifique unívocamente.
- La cantidad de hogares en cada unidad censada se deben poder calcular.
- Debemos considerar información incompleta ya que cuando se realiza un censo no siempre se logran relevar todos los datos, puede ocurrir que una persona no se encuentre cuando pasó el censista o simplemente no quiso completar en su totalidad los formularios del censo. Para poder considerar esta incompletitud clasificamos a los hogares según el tipo de información que sus ocupantes provean, aun cuando sea nula, esto es necesario debido a que en regiones pequeñas si los hogares relevados son insuficientes y además de eso se releva menos información estaremos aumentando significativamente el riesgo de divulgación para los individuos o generando estadísticas erróneas. Además para calcular la cantidad de personas en un hogar debemos tener en cuenta que las personas ausentes afectan la calidad de la información, por esto se tipifica a las personas con un estado.
- Se tienen que considerar las características definidas para las personas mencionadas en el capítulo III y teniendo en cuenta los límites que impone la ley sobre ellos, en síntesis de una persona debemos conocer:
  - Estudios.
  - Familia.
  - Datos personales.
  - Ocupación.
  - Donde vive y con quien
  - Si su vivienda es propia o alquilada y si cuenta con las necesidades básicas.

- Se deben poder seleccionar los filtros a utilizar sobre los radios censales cuando se calcule un Indicador Demográfico y si este implica operaciones entre dos o mas geografías se deben devolver los resultados de las nuevas regiones generadas por el solapamiento de las mismas, es decir de sus intersecciones.
- Para poder comparar dos geografías es necesario tener estadísticas precisas acerca de las áreas en cada una de ellas, en general las estadísticas deben ser generadas a partir de los datos originales, pero de esto solo consideraremos que son ingresados.

Se consideran que los valores que se utilizan para caracterizar el cálculo del Indicador en base a la información, sean parametrizados para proporcionarle mas flexibilidad al modelo.

## FUNDAMENTACIÓN DEL PARADIGMA

### Modelo de Datos

Cada modelo tiene semántica y notación, que pueden adoptar varios formatos que incluyen texto y gráficos. Rumbaugh [13] afirman que un propósito fundamental de los modelos es que permiten “captar y enumerar exhaustivamente los requisitos y el dominio del conocimiento, de forma que todos los implicados puedan entenderlos y estar de acuerdo con ellos”.

Nuestro modelo describe la aplicación usando el paradigma orientado a objetos, fundamentamos la elección debido a que un modelo que aplica este paradigma permite la creación de bases de datos orientadas a la información geográfica (Geodatabase).

Geodatabase define un modelo genérico para información geográfica, el cual puede ser usado para definir y trabajar con una amplia variedad de usuarios o aplicaciones específicas, implementando diferentes comportamientos permitiendo modelar y manejar entidades que poseen atributos espaciales (i.e., localización, forma y topología) y otros atributos no vinculados con el espacio, denominados atributos no espaciales (p.ej., la capacidad, nombre). Esta separación requiere que los usuarios de un GIS se familiaricen con dos notaciones diferentes: una para manipular los datos espaciales y otra para los no espaciales.

Recientemente, han surgido nuevos modelos de datos espaciales basados en la Orientación a Objetos que resuelven, en parte, los problemas semánticos antes descritos ([20] y [21]). La principal cualidad de los modelos espaciales orientados a objetos es la naturalidad con que se modela el problema. Cada entidad espacial o no-espacial del dominio de aplicación del SIG se representa mediante un objeto, el cual captura las propiedades y comportamiento de la entidad representada.

El modelado orientado a objetos de un SIG consiste, esencialmente, en describir las clases de objetos espaciales y no espaciales relevantes al dominio de aplicación del SIG, así como establecer las relaciones entre estas clases. A diferencia de las clases de objetos no-espaciales, las clases espaciales guardan entre sí muchos elementos estructurales y de comportamiento comunes, que se repiten de una clase a otra y de una aplicación a otra.

Estos aspectos y soluciones motivan el desarrollo de este trabajo, cuyo objetivo fue el desarrollo de una arquitectura que pudiera ser aplicada como patrón para solucionar problemas censales.

### *Lenguaje de Modelado*

Para documentar el análisis del modelo utilizamos UML (Unified Modeling Language) el cual es un lenguaje que permite modelar, construir y documentar los elementos que forman un sistema software orientado a objetos. [13]

El proceso a seguir en un desarrollo orientado a Objetos (OO) es complejo y esta formado por una serie de actividades y subactividades. En este trabajo abarcamos la fase de planificación y especificación de requisitos, la cual se corresponde con un borrador del modelo conceptual y con una definición de casos de uso de alto nivel.

Los casos de uso no son mas que la descripción de los eventos de un actor que usa un sistema para completar un proceso [14], modelan la funcionalidad del sistema, expresando la transacción entre actores y sistema.

En nuestro trabajo los casos de uso más críticos, importantes y que conllevan un mayor riesgo, se amplían. Dejando la definición en formato expandido del resto de casos de uso para cuando sean tratados en posteriores ciclos de desarrollo. Dejamos la referencia [14] –[26] para quienes deseen ampliar sobre este tema.

### *Representación de las relaciones:*

UML define la forma en la que se representa la relación que existe entre dos clases y su propósito dentro del modelo.

Distinguiremos las siguientes asociaciones

#### Relación Directa:

Representa que una variable de instancia mantiene una referencia de otro objeto.





Agregación:

Tipo especial de relación entre clases de entidades que representa una conexión “es parte de” o “tiene un”.

“el todo” ◇———— “la parte”

*Representación*

Composición:

Es un tipo especial de agregación en las que hay una fuerte pertenencia y la existencia de las ocurrencias de las clases de entidades vinculados coincide.

“el todo” ◆———— “la parte”

*Representación*

Herencia:

Mecanismo mediante el cual el tipo de una clase puede ser definido sobre la base de la definición del tipo de otra clase. Gracias al mecanismo de herencia, la subclase hereda los atributos que definen la estructura de la superclase y los métodos que caracterizan su comportamiento. Además, la subclase puede añadir nuevos atributos y métodos para completar su definición. Las relaciones de herencia organizan las clases en un tipo de grafo especial denominado jerarquía de herencia o árbol de herencia.

◀————

*Representación*

Multiplicidad:

Sobre las asociaciones se definen restricciones que se observan en el problema y que controlan si las ocurrencias de las clases de entidades que intervienen en el tipo de relación pueden o no ser múltiples conexiones con ocurrencias de las demás clases. Representamos la multiplicidad en los extremos de la asociación, estas expresiones son simplemente números, rangos o combinación de ambos.

Patrones

Explicamos muy brevemente, la noción de patrón de diseño empleada en este trabajo, ya que es un tema muy amplio del que solo nos interesa a modo de guía para entender su propósito.

Un patrón de diseño de software es una solución reutilizable a un problema recurrente en el desarrollo de software [15], [16]. Su reutilización reduce el esfuerzo, tiempo y costos empleados durante el diseño de software [17]. Los patrones de diseño capturan la estructura

y dinámica de una solución que se repite múltiples veces durante el desarrollo de diferentes aplicaciones, generalmente, en un contexto o dominio determinado [18], [19].

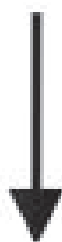
Un patrón de diseño está compuesto por un conjunto de clases que, mediante la delegación y herencia, proporcionan una solución robusta y modificable a un problema común o recurrente. Estas clases son reutilizadas en el diseño de nuevas aplicaciones; para lo cual se requiere su adaptación o refinamiento en base a los requerimientos de la aplicación [17].

En el siguiente capítulo utilizamos estos conceptos, generamos conclusiones en base a los capítulos anteriores sobre el problema planteado, y sintetizamos las etapas que dieron como resultado nuestro modelo.



# CAPÍTULO V

## EL MODELO





## INTRODUCCIÓN

Iniciamos la etapa de documentación, en ella sintetizamos el análisis realizado para obtener el modelo en base a todo lo relevado.

Como mencionamos en el capítulo II, un SIG trata aspectos complejos como la representación de relaciones espaciales entre objetos del mundo real y sus características geográficas, permitiendo manipular datos espaciales y no espaciales, analizamos las características espaciales y no espaciales por separado, derivando en un modelo geográfico y un modelo conceptual respectivamente.

Para llegar al modelo conceptual, identificamos los casos de uso generando los diagramas de contexto necesarios, sobre los cuales mostramos los conceptos relevantes a nuestro modelo. Nos concentramos en hallar características y organizar la información que debe ingresar a nuestro modelo, de tal forma que las preguntas que se hagan puedan resolverse.

En el modelo geográfico, identificamos las entidades del modelo conceptual sobre las que necesitamos representar atributos espaciales y definimos comportamiento, ubicación con respecto a uno o a muchos sistemas de proyección y geometría.

Entonces organizamos este capítulo en 3 secciones

Análisis de Datos no espaciales

Análisis de Datos espaciales

Integración de los datos espaciales – no espaciales

### Análisis de Datos No Espaciales

A continuación definimos los casos de uso basándonos en el ejemplo de aplicación de nuestro modelo descrito en el capítulo IV.

## CASOS DE USO

### Casos de Uso de Alto Nivel

La figura 5.1 muestra los casos de uso que hemos identificado.

Caso de Uso: Seleccionar Censos

Actores: Persona

Propósito: Capturar la población exacta de las geografías a comparar.

Resumen: La persona selecciona los años de los censos obteniendo características espaciales y no espaciales.

Caso de Uso: Consultar datos

Actores: Persona

Propósito: Obtener datos poblacionales para diferentes indicadores demográficos

Resumen: Selecciona la consulta y los parámetros necesarios para realizarla.

Caso de Uso: Incorporar información censal

Actores: Administrador del Sistema

Propósito: Incorporar la información de un nuevo censo a la Base de Datos

Resumen: Deberá ingresar los datos relevados en el censo, espaciales y no espaciales.

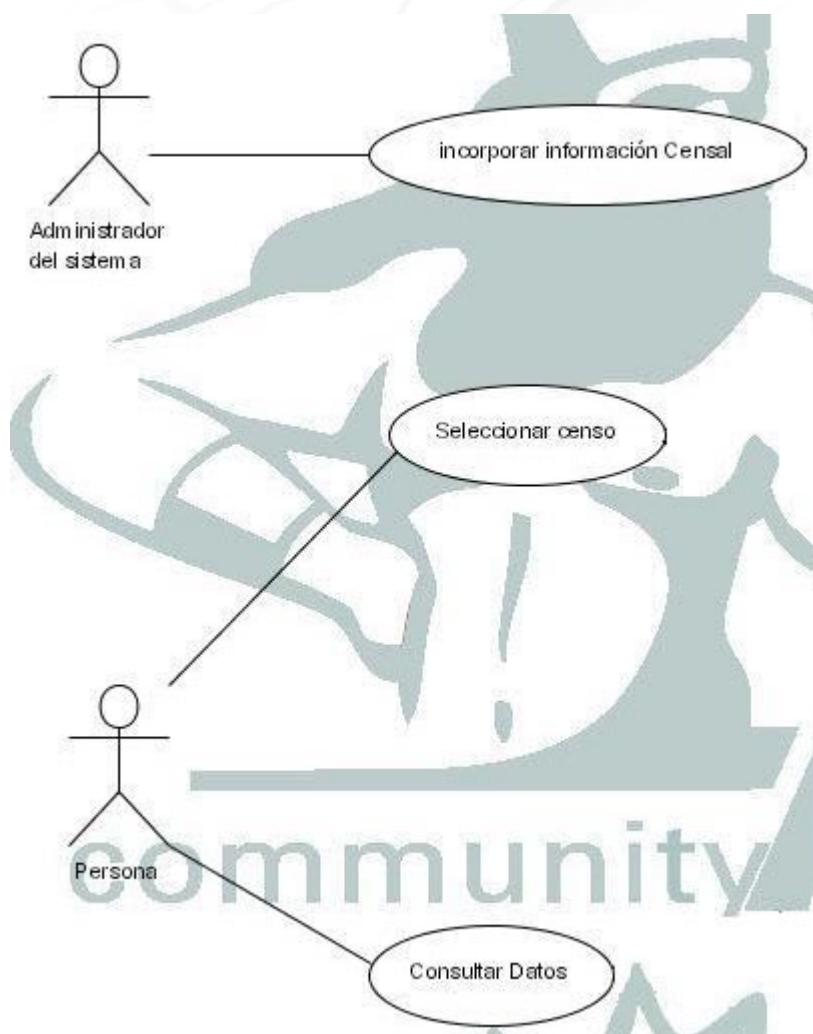


Fig 5.1 Modelo de contexto. Diagrama UML diseñado en Poseydon CE2.

## Caso de Uso expandido

Debido a la complejidad de los casos de uso definidos en la figura 5.1, se presenta el caso de uso expandido descrito en la Figura 5.2. De esta manera reducimos la complejidad y ampliamos la información que facilitara el diseño del modelo conceptual.

Caso de Uso: Incorporar información censal

Caso de Uso expandido : Incorporar datos relevados

Caso de Uso expandido : Incorporar año

Caso de Uso expandido : Incorporar región

Caso de Uso: Seleccionar Censos

Caso de Uso expandido : Seleccionar región Geográfica

Al seleccionar una región geográfica se podrán seleccionar regiones, fracciones o radios censales definidos para el censo, pudiendo seleccionar varios.

Caso de Uso expandido : Seleccionar Año

Caso de Uso: Consultar Datos

Caso de Uso expandido : Introduce parámetros

Caso de Uso expandido : Ejecuta consulta

Caso de Uso expandido : Superponer zonas de cada censos

Caso de Uso expandido : Acotar a la región seleccionada

Caso de Uso expandido : generar información.

## Fase de Construcción del Modelo Conceptual

### *Análisis*

En la fase de análisis investigamos los conceptos relacionados con el subconjunto de casos de uso que definimos en la etapa anterior. En particular, profundizamos en los datos que se pueden ingresar como parámetros para las consultas que se soliciten.

Mostramos de manera simple la solución, sin detallar su implementación.

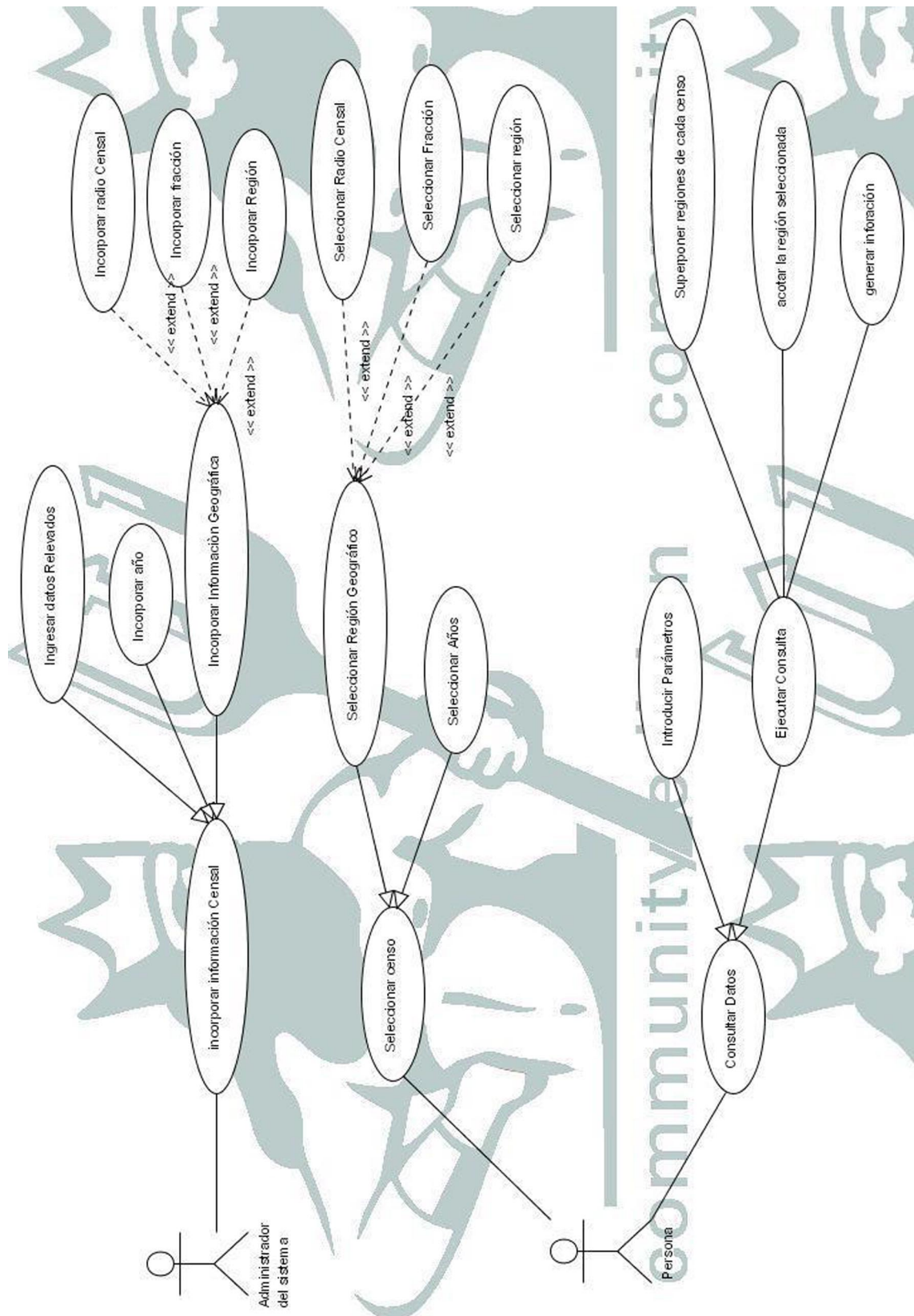


Fig 5.2 Modelo del contexto del sistema. Diagrama UML diseñado en Poseydon CE2.

Como primer paso identificamos los conceptos que tomamos en cuenta al representar gráficamente el modelo. Para representar estos conceptos se va a usar un diagrama de estructuras estáticas de UML.

Seguimos los siguientes pasos:

Lista de conceptos candidatos:

Definimos estos conceptos en base a la información que fue relevada en el capítulo III, cuando investigamos los censos, en particular tomamos aquellas características que se refieren a censos poblacionales. Los conceptos a tener en cuenta son los siguientes:

- Vivienda Colectiva
- Vivienda Familiar
- Persona Censada
- Persona Ausente
- Hogar
- Familia
- Vinculo
- Edificio
- Electrodomésticos
- Empresa
- Obra social
- Estudios
- Institución
- Servicio
- Trabajo
- Censo
- Región
- Radio Censal
- Fracción
- Lote
- Manzana
- Indicador demográfico
- Filtro

De esta manera obtenemos un modelo conceptual en base a los conceptos antes mencionados sin incluir atributos ni asociaciones

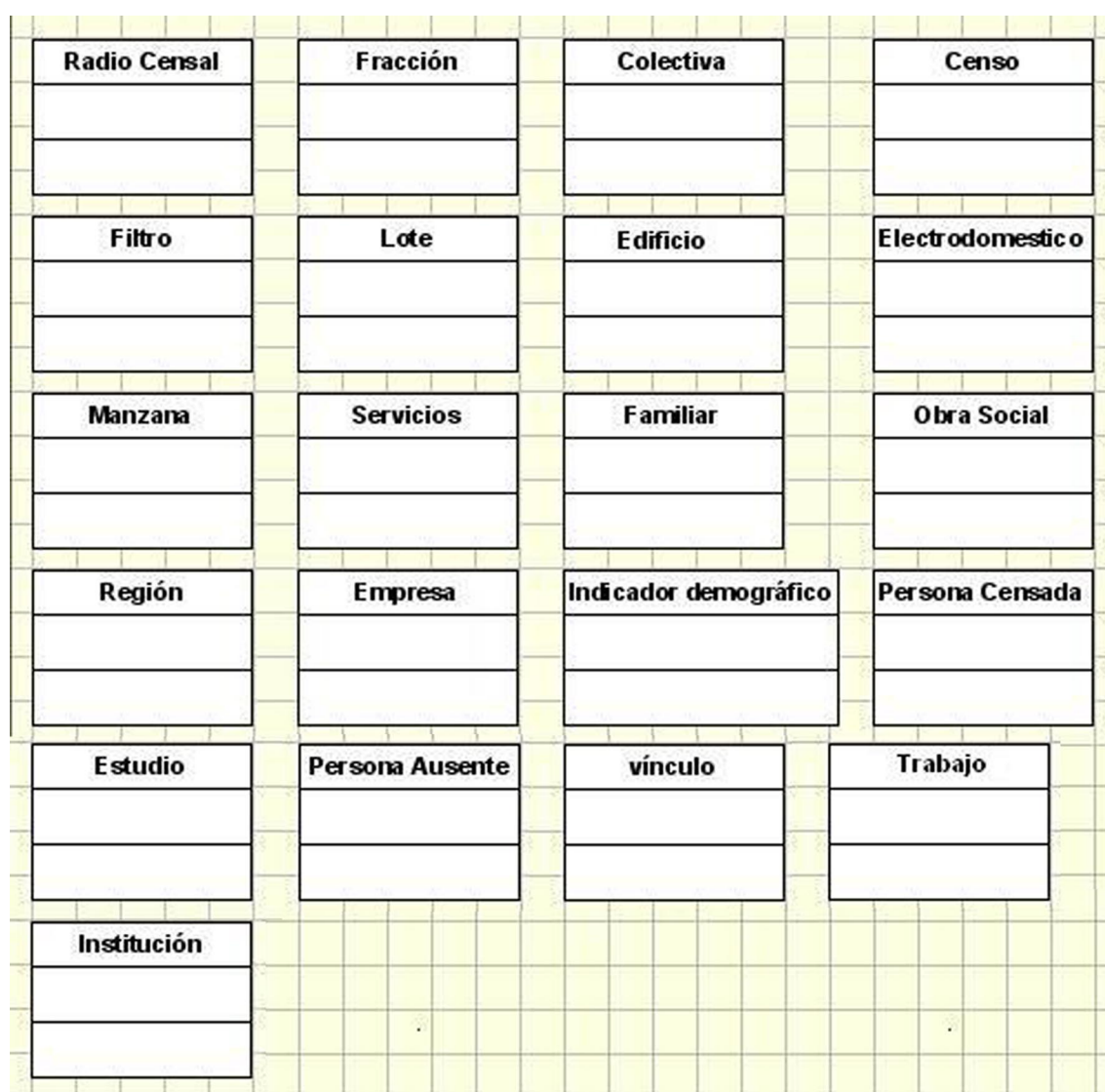


Figura 5.3 Diagrama Conceptual Inicial. Diagrama UML diseñado en Poseidón CE2.

## Definición de Asociaciones y Multiplicidad

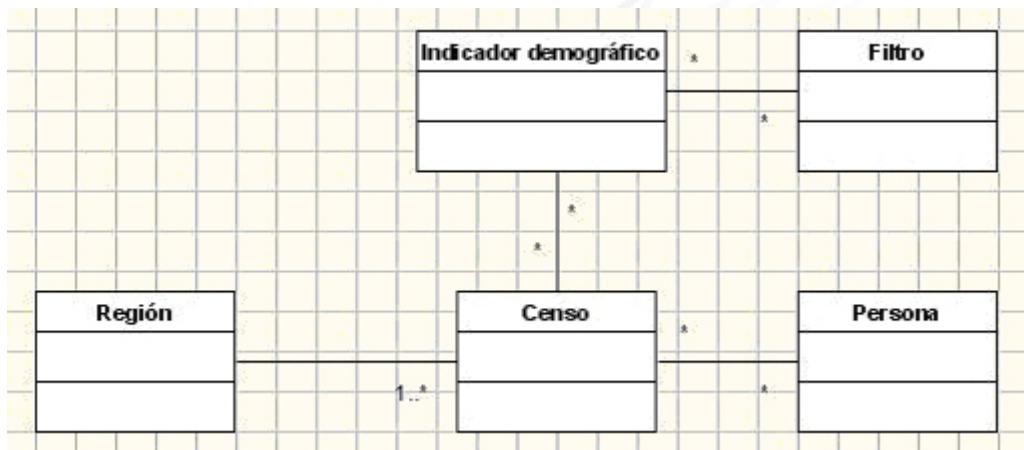
Luego, añadimos las asociaciones necesarias para ilustrar las relaciones entre los conceptos. Para realizarlo utilizamos las definiciones mencionadas en el capítulo IV, en las que explicamos cómo se representan en UML las asociaciones y la multiplicidad entre los conceptos.

Tomamos grupos de conceptos y al final mostramos el modelo conceptual resultante de la unión de ellos, optamos por realizarlo de esta manera para poder simplificar la comprensión del mismo.



*Censo, región, indicador demográfico, filtro, persona*

El censo se relaciona con la región sobre la cual se realizó y con las personas que conforman la población censada. También se relaciona con el Indicador Demográfico, el cual requiere cierta información, nos referimos al mecanismo que utilizará para resolver su fórmula, de aquí surge la relación con otra entidad a la cual llamamos Filtro, como explicamos en el capítulo IV un filtro posible es “Redondeo”, si este filtro ingresa como parámetro para el indicador, cuando se resuelva la fórmula se redondea el resultado. La elección del filtro que se debe utilizar esta a cargo del investigador que requiere la consulta. El esquema resultante de las asociaciones que planteamos es el siguiente:



*Figura 5.4 El Censo. Diagrama UML diseñado en Poseidón CE2*

*Región, radio censal, fracción, lote, manzana*

Sobre los conceptos de Región – Radio Censal – Fracción - Lote - Manzana definimos la relación de composición, ya que como se mencionó en el capítulo IV la región a censar se subdivide en fracciones, estas en radios censales, los mismos en manzanas y por último las manzanas en lotes. La multiplicidad también fue determinada por la información obtenida de la metodología que se asigna a los censistas para relevar la información en un censo. El radio censal asignado a un censista no puede tener menos de 3 manzanas. En la figura 5.5 se puede observar que se indica la cardinalidad en muchos, con el propósito de que se adecue a las limitaciones del censo que se este realizando.

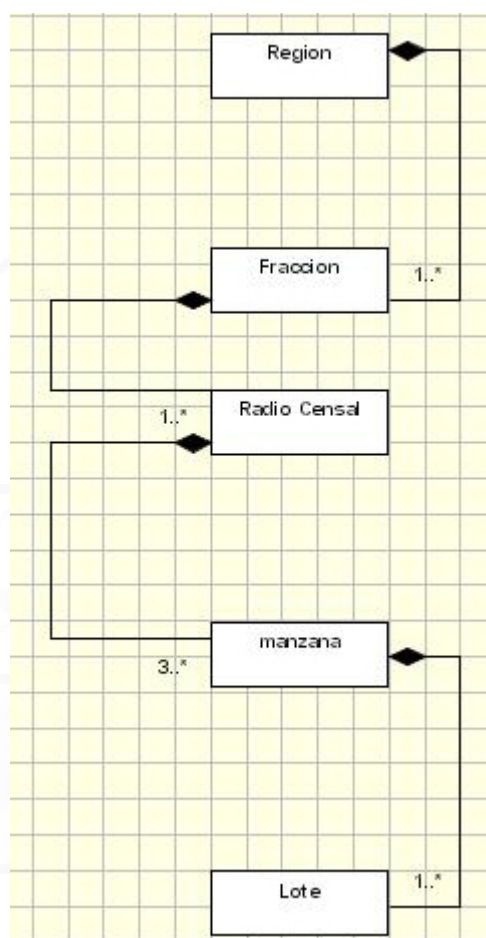


Figura 5.5 División de la Región a censar. Diagrama UML diseñado en Poseidón CE2.

Vivienda colectiva, vivienda familiar, hogar, alojamiento, local, edificio, complejo de edificios, electrodoméstico y servicio

Sobre la vivienda se plantea una jerarquía, ya que la vivienda puede ser colectiva o familiar, a su vez existe una relación de composición entre ellas debido a que una vivienda colectiva puede estar conformada por viviendas familiares, tal y como se describió en el capítulo III. El Alojamiento, Hogar, edificio, local, complejo de edificios se contemplan en esta estructura ya que una vivienda colectiva puede ser un hotel, convento, pensión, etc. La vivienda describe las comodidades que esta ofrece a sus habitantes, es por esto que son importantes para medir sus condiciones de vida de la población, indicando tipo de construcción, cantidad de ambientes, electrodomésticos y servicios con los que cuenta. También se le adicionaron atributos acerca de su construcción.



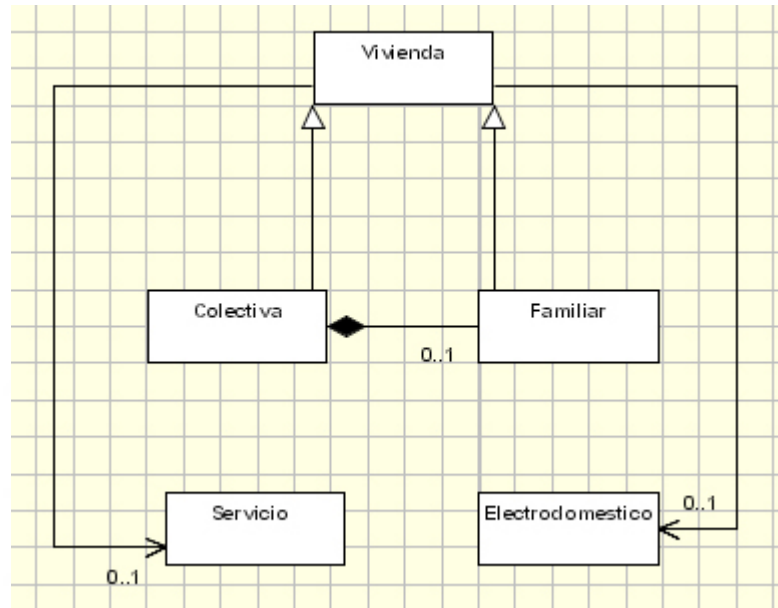


Figura 5.6 La vivienda. Diagrama UML diseñado en Poseidón CE2.

#### Persona, obra social, estudio, institución, empresa y vínculo

Persona es la entidad sobre la que se centra toda esta investigación, nuestro modelo debe soportar toda la información referente a la misma. Modelamos a la persona censada y a la persona ausente, ya que ambos tipos de personas, tienen un efecto importante a la hora de calcular un indicador demográfico.

Utilizamos el patrón estado (STATE) para representar si la persona estuvo ausente o no en el momento en el que el censista visitó la vivienda, este patrón permite a un objeto cambiar su comportamiento cuando su estado interno cambia, es decir que según el estado que tienen deben comportarse de manera distinta.[16]

Nos parece importante resaltarlos porque cuando se calculan los indicadores demográficos que incluyan personas, la persona presente se comportará distinto que la ausente.

También representamos el vínculo que existe entre las personas y si éstas conviven o son familiares. De una persona necesitamos, también modelar si posee ó no obra social, sus estudios y en que tipo de institución los realizó, también es importante representar si trabaja y su lugar de trabajo.

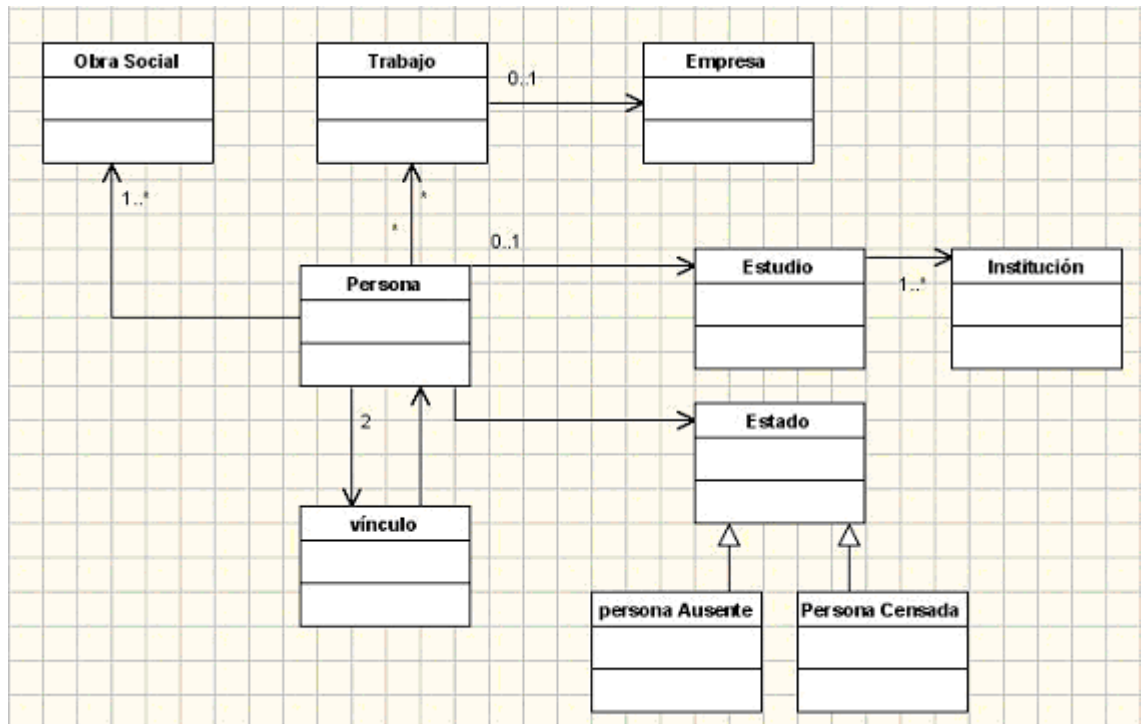


Figura 5.7 Entidad persona y sus relaciones. Diagrama UML diseñado en Poseidón CE2.

## Definición de Atributos

Añadimos los atributos necesarios para contener toda la información que se necesite conocer de cada clase.

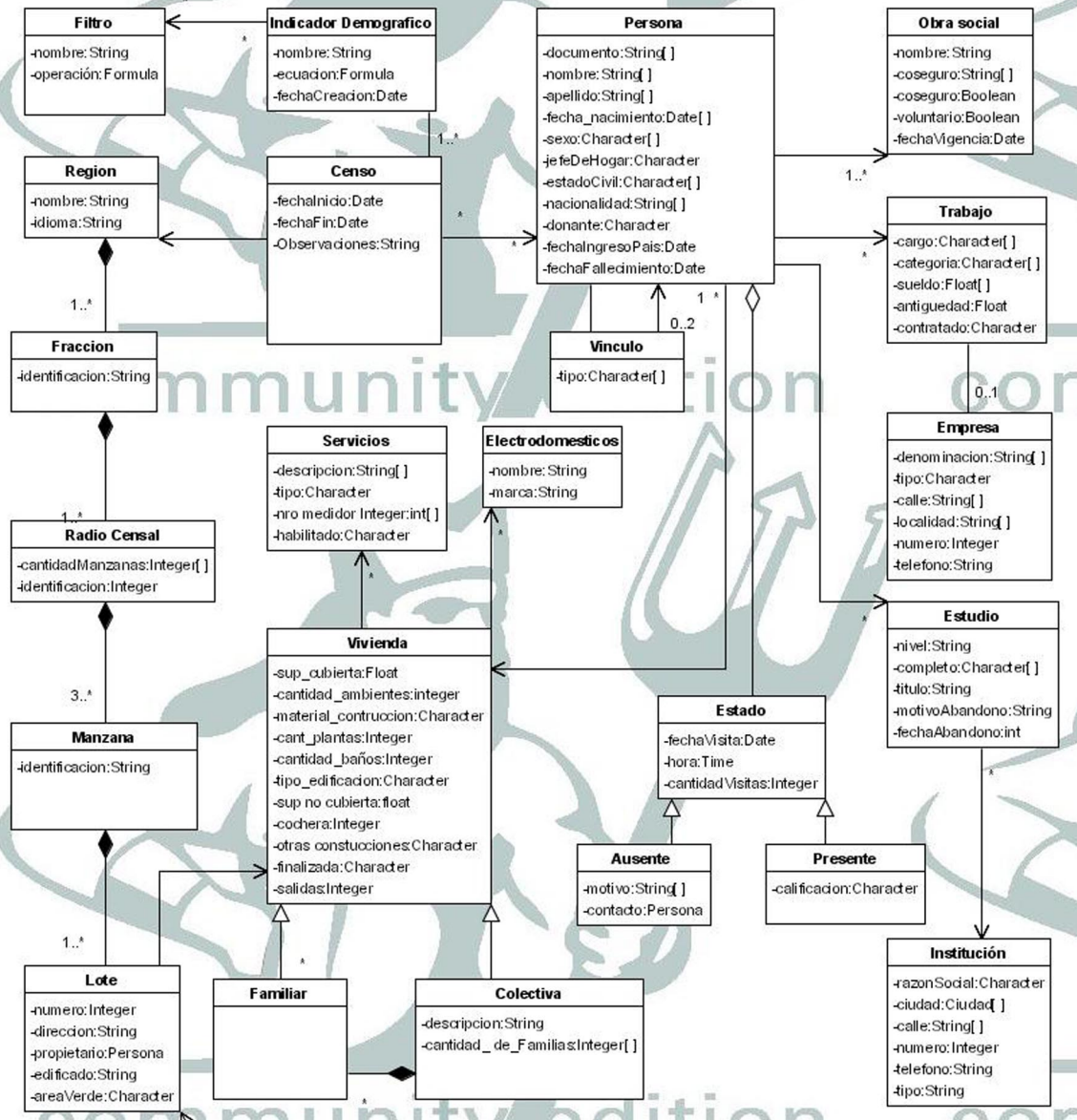


Figura 5.8 Modelo conceptual con atributos. Diagrama UML diseñado en Poseidón CE2.

## Definición del Comportamiento

Para hacer mas claro el propósito de alguna de las clases, agregamos comportamiento en la figura 5.9 mostramos el Modelo conceptual final con el comportamiento mas relevante para los nuestros fines.



Todos los atributos y el comportamiento definidos en las clases se determinaron teniendo en cuenta las definiciones del capítulo III, en el cual explicamos los conceptos que se derivan de un censo.

La definición formulada hasta ahora constituye la base conceptual sobre la cual se fundamentó la formulación de los patrones de diseño, que se usaron para el diseño final del Modelo, así como también para el modelo geográfico que planteamos en la siguiente sección.

## **Análisis de Datos Espaciales**

Como mencionamos en el capítulo II, el mapa es probablemente la forma mas familiar que conocemos para representar datos geográficos, ya que referencia en dos dimensiones un SIG, también quedó claro que a medida que se le va adicionando información y éste va creciendo, manipularlo manualmente es prácticamente imposible. En esta sección modelamos características geográficas para las clases de nuestro modelo conceptual que lo necesiten, teniendo en cuenta, qué aspectos comunes debemos representar. Según lo analizado en el capítulo II podemos decir que vamos a tener que encontrar la forma de representar la geometría de las entidades espaciales, la cual está determinada básicamente por cuatro atributos:

### Forma:

Puede ser puntual, lineal o poligonal dependiendo de la escala utilizada.

### Localización, dirección y orientación:

Cuando la forma del objeto es georeferenciada, su referencia determina, también, la ubicación del objeto, su dirección y su orientación.

La figura 5.10 muestra tres clases que representan la estructura y comportamiento de tres tipos de entidades espaciales con formas puntual, lineal y poligonal.

En ellas se puede observar que el nombre, la escala, el sistema de geo-referencia y la geometría son atributos comunes a todas los tipos de entidades espaciales independientemente de su forma.

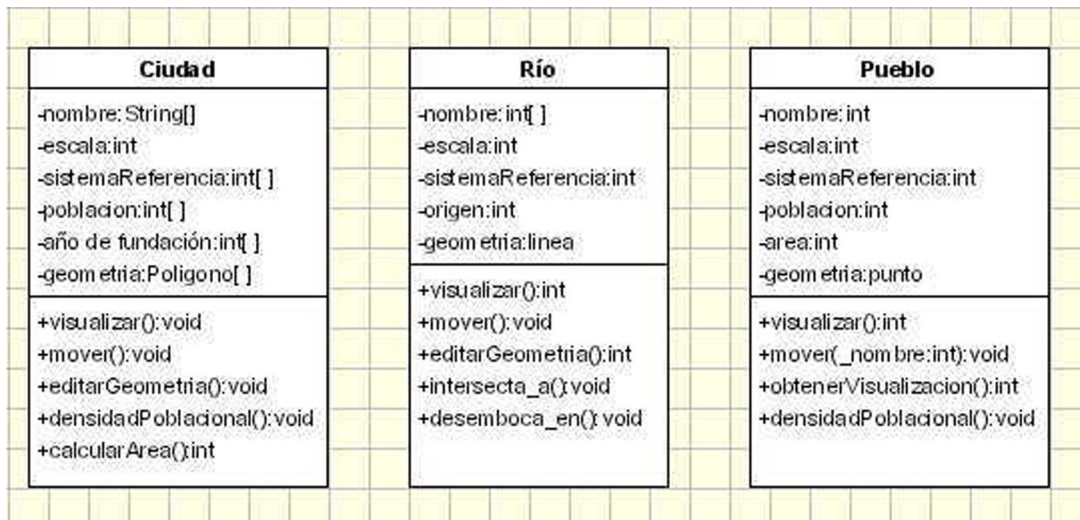


Figura 5.10. Clases espaciales con diferente geometría. Diagrama UML diseñado en Poseidón CE2.

Teniendo en cuenta lo anteriormente explicado pasamos a mostrar como aplicamos las características mencionadas en nuestro modelo, tomamos la clase conceptual “LOTE”, le adicionamos características geográficas y la utilizamos para explicar los conceptos que definimos a continuación, primero definimos al objeto espacial, luego vemos como representar su ubicación y su sistema de proyección, así como también su geometría, ya que lo que necesitamos es representar entidades espaciales que existen en un espacio cuya geometría está determinada por uno o varios sistemas de proyección.

Entonces al representar a un objeto espacial estamos definiendo la estructura y el comportamiento que tiene una entidad del mundo real en un espacio dado.

Para la clase lote consideramos la siguiente estructura:

- Un identificador el cual nos permitirá identificar de forma única al Lote.
- Su geometría en adelante **geo** que será el conjunto de puntos que lo representan, cada punto esta representado por un par ordenado (x,y).
- Sus atributos espaciales y no espaciales a1,a2,..an de los cuales representamos las propiedades mas relevantes para nuestro modelo.

Añadimos las características geográficas mencionadas a la clase conceptual Lote, en adelante **GeoLote**. Ver figura 5.11

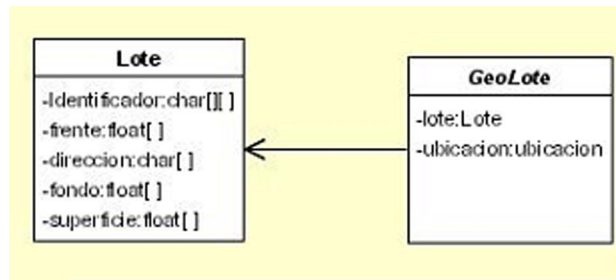


Figura 5.11 Separación de las características geográficas del Lote en una nueva clase.  
Diagrama UML diseñado en Poseidón CE2.

A continuación explicamos brevemente el funcionamiento para la clase Lote, redefinimos la clase GeoLote y para hacerlo aplicamos el patrón Decorador. Decorador es un patrón de diseño, su intención es agregar responsabilidades a un objeto dinámicamente para que provea una alternativa flexible a la creación de subclases y así extender la funcionalidad de una clase. La diferencia entre decorar y crear subclases, es que la primera es dinámica y la segunda es estática, esto es justamente lo que buscamos, al usar decoradores podemos agregar funcionalidades a algunos objetos sin redefinir la jerarquía conceptual, un decorador replica la interface del objeto decorado; también define comportamientos adicionales, pasa los pedidos al componente y puede realizar acciones adicionales antes o después de pasarlos, para ampliar sobre los fundamentos de utilizar este patrón de diseño para modelar características espaciales vease [16].

De esta manera podemos definir objetos y decorarlos con diferentes comportamientos, por ejemplo comportamiento espacial y dirigirnos al objeto original o a cualquiera de sus decoradores. Los decoradores imitan el protocolo del objeto conceptual y delegan la implementación del protocolo no espacial a ese objeto. Por lo tanto el Lote tendrá mas posibilidades de resolver cuestiones geográficas con estas nuevas características.

Aplicar este patrón tal y como se explica en [30] [22] nos llevo a definir la clase geoLote de forma tal que las instancias trabajen decorando a los objetos lotes. La clase GeoLote define el comportamiento geográfico que necesitamos en nuestro Modelo, por ejemplo incluye un método que calcula el ancho de un lote. Dado que hay una relación entre el decorador GeoLote y el Lote, el primero puede delegar al objeto conceptual la ejecución del comportamiento ya definido.

## Integración de los datos Espaciales – No Espaciales

Sobre los atributos no-espaciales, definimos una clase en el modelo conceptual que define las características no espaciales del Lote. Por ejemplo, dirección. Ver figura 5.9

Los atributos espaciales pueden ser geométricos o espaciales, los **geométricos** nos permitirán representar la geometría del Lote, mediante la localización, forma, orientación



y dirección que tiene para una escala dada.

Para representar la geometría vamos a modelizar los puntos, líneas, polígonos cuyas coordenadas vienen dadas por un sistema de proyección determinado.

Entonces la geometría va a ser una clase abstracta y debemos generalizar la estructura y comportamiento común de las clases Punto, Línea y Polígono, las cuales definen la geometría de las instancias de **GeoLote**. Ver figura 5.12

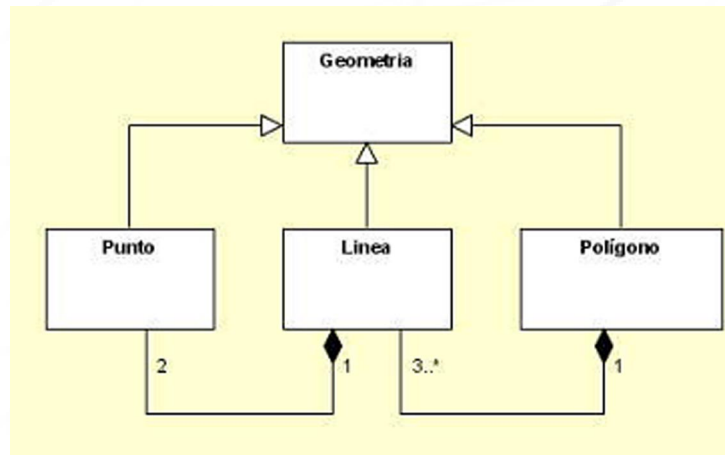


Figura 5.12 Clase geometría. Diagrama UML diseñado en Poseidón CE2.

La clase punto define como atributos las coordenadas x, y, z, que están geo-referenciadas por el sistema indicado por el atributo sistema Referencia en la clase **GeoLote**. Cada instancia de Línea tiene asociado un conjunto ordenado de instancias de Punto que determinan su forma. Similarmente, cada instancia de Polígono tiene asociado un conjunto ordenado y secuencial de líneas que determinan el borde del polígono. Ver figura 5.13

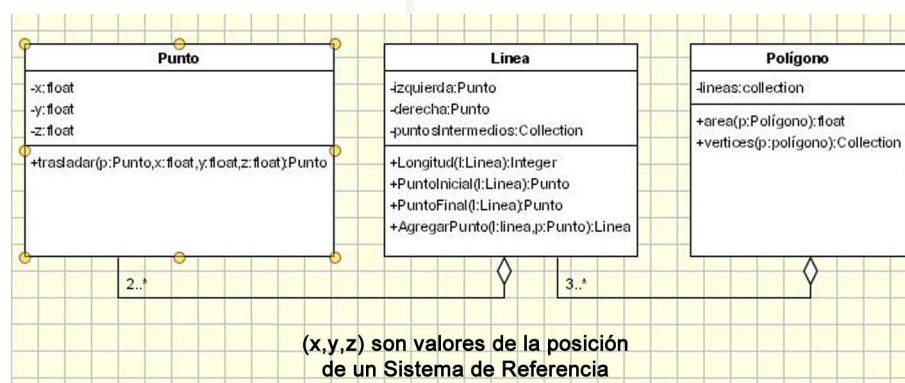


Figura 5.13 Punto, Línea y Polígono. Diagrama UML diseñado en Poseidón CE2.

Los atributos izquierda y derecha de la clase Línea son usados para establecer las relaciones topológicas de las instancias de Línea y Polígono.

Para representar las **relaciones espaciales** tomamos las relaciones que se dan en el espacio entre el Lote y otras entidades espaciales.

Entre las relaciones espaciales existentes, consideramos las de orden espacial, topológico (ver figura 5.14) , el modelo propuesto se puede ampliar agregando más relaciones espaciales, según se requieran, por ejemplo se podrían considerar también las relaciones difusas, sobre las que encontramos un artículo [23] que se da como referencia para aquel que desee profundizar en el tema.

De orden espacial	Topológicas
A está frente a B	A intersecta a B
A esta detrás de B	A es adyacente a B
A está a la derecha de B	A está dentro de B
	A se solapa con B
	A es cubierta por B

Figura 5.14 Ejemplos de Relaciones Topológicas y Espaciales para dos entidades A y B cualesquiera.

La estructura estandar para definir objetos geográficos, es la que se muestra en la figura 5.15 que mostramos a continuación, en ella representamos lo enunciado hasta ahora.

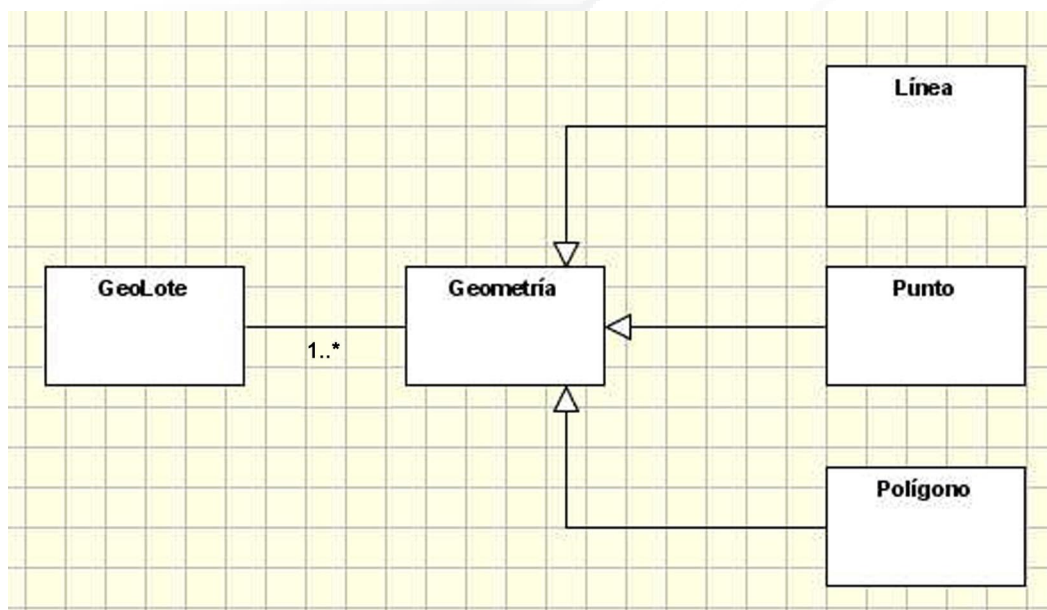


Figura 5.15 Estructura definición espacial. Diagrama UML diseñado en Poseidón CE2.

Aún habiendo llegado a definir la geometría y el objeto espacial GeoLote necesitamos representar otra característica más, ya que según lo analizado hasta ahora, la representación del lote puede variar de acuerdo al sistema de proyección que se utilice, es decir si cambia el sistema de proyección, debemos crear otro objeto espacial lote con otra escala definida.

Los decoradores, son una solución mas dinámica que la creación de subclases ya que

las responsabilidades se pueden agregar o quitar en tiempo de ejecución al adjuntar y quitar componentes al objeto decorado, con ellos la aplicación resultante manipula dos tipos de objetos los conceptuales y los geográficos definiendo dos niveles separados.

De esta manera el modelo permanece flexible porque podemos definir múltiples decoradores para un objeto específico e incluso decorar a un decorador anidando así especificaciones geográficas. Cuando estamos ampliando aplicaciones existentes podemos hacerlo de manera transparente porque el modelo conceptual no sufre modificaciones por lo tanto la aplicación original aún puede ser usada como fue concebida originalmente.[22]

Aplicando este método los objetos geográficos siempre saben su ubicación y tienen una geometría y la ubicación contendrá la posición del objeto georeferenciado y algún comportamiento que es usado con datos temporales, por ejemplo datos que cambian su posición en función del tiempo.

Además plantea representar para la ubicación un sistema de proyección para mantener información acerca de los sistemas que se utilicen. Para realizarlo define el Patrón Sistema de Referencia que también aplicamos para poder desagregar del objeto espacial su sistema de proyección.

El patrón definido en [30] [22] que aplicamos, define que cuando necesitamos combinar información de diferentes objetos, debemos asegurarnos que todos ellos estén utilizando el mismo sistema de proyección, de otra manera necesitamos realizar conversiones, lo mismo ocurre cuando el sistema de proyección que utilizamos para presentar información al usuario es distinto al sistema de proyección utilizado al ingresar los datos. Este Patrón ayuda entonces a definir el contexto de una ubicación, provee un juego de operaciones legales que son aplicables a la aritmética del sistema de proyección correspondiente incluyendo operaciones relativas al tiempo.

Desde el punto de vista de la implementación la aplicación nos provee operaciones para cambiar dinámicamente de un sistema de referencias a otro.

Desde el punto de vista de diseño nos ayuda a expresar esta decisión de diseño de forma clara y sin ambigüedades.

Sin la información que el sistema de proyección nos provee, el juego de puntos que representan la ubicación de objeto geográfico no tiene sentido, trabajar con varios sistemas de referencia también es complejo, ya que aun cuando es posible implementar algoritmos para traducir ubicaciones de un sistema de referencias a otro, las semánticas de estos cambios no son explícitamente acordadas durante el diseño.

Se han separado la ubicación de sistema de referencia para permitir configurar dinámicamente al objeto ubicación con diferentes sistemas de referencias durante el tiempo

de vida de un objeto geográfico. Claramente esta solución de diseño es mejor que crear subclases de ubicación de acuerdo a diferentes sistemas de referencia ya que el diseño resultante sería más rígido y no permitiría cambiar el sistema de referencias dinámicamente.

*La clase ubicación* implementa el comportamiento básico para permitir la representación de las coordenadas geográficas dentro de un sistema de referencia.

*La clase sistema de referencia* define un protocolo abstracto que se usa para describir el contexto en el que se define una ubicación también define un juego de operaciones legales en este contexto, en otras palabras cada instancia de la clase sistema de referencia describe como se comportan las medidas en la ubicación definida.

*La clase Medida* describe las unidades de los valores de la clase ubicación [25] se define un objeto (ubicación), que encapsula la relación entre una medida y un sistema de referencia, la clase ubicación conoce el valor de su posición y el sistema de referencia utilizado. Este patrón permite gran flexibilidad ya que soporta varias maneras de representar entidades del mundo real basada en diferentes abstracciones terrestres. Más aún permite cambiar o traducir ubicaciones de un sistema a otro sin afectar al objeto referenciado.

Cada sistema de proyección implementa un conjunto de operaciones reales, tales como el cálculo de distancias entre puntos, la comparación de elementos y el cálculo de áreas, también especifica operaciones de traslación a otros sistemas de proyección.

Las distancias, las áreas y cualquier otra operación que dependa de la posición del objeto geográfico se expresan mediante colaboraciones entre la ubicación de ese objeto y el sistema de proyección de ubicaciones.

También debemos considerar que pueden aparecer objetos puramente geográficos que no representan una vista de ningún objeto conceptual sino que se caracterizan por su relevancia geográfica, en otras palabras estos objetos tienen las mismas características espaciales que un decorador pero no están relacionados con ningún objeto conceptual, basándonos en las similitudes entre los decoradores que agregan información espacial a los objetos conceptuales y los objetos puramente geográficos se define una clase abstracta que agrupa el comportamiento común de estos objetos puramente geográficos además de los encapsuladores del modelo conceptual.[22]

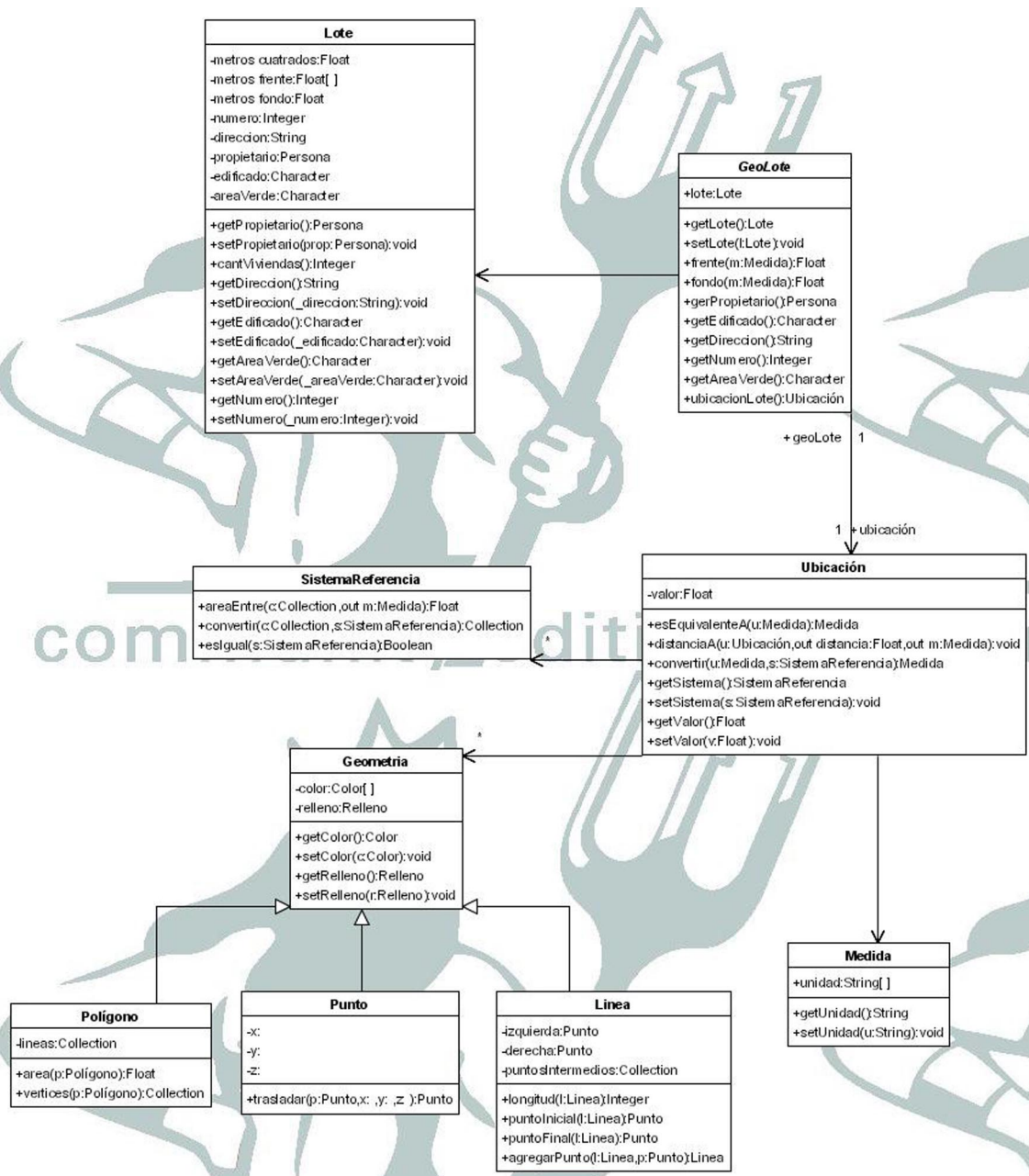


Figura 5.16 Redefinición de las características geográficas de la clase lote aplicando [50].  
Diagrama UML diseñado en Poseidón CE2.

La figura 5.16 muestra el resultado de aplicar los patrones definidos en [22] a la clase lote. Lo que debemos tener en cuenta es que cuando agregamos o quitamos una característica conceptual debemos agregar o quitar la característica correspondiente en el decorador.

## El Modelo: Modelo Conceptual - Modelo Geográfico

Una vez definidas las características geográficas, el modelo resultante completa la funcionalidad deseada, aplicando el método tomado de [22], para completar la dinámica necesaria en cuanto a la representación del Lote. En la figura 5.17 mostramos el modelo obtenido.

La principal aplicación es en el modelado de censos poblacionales. Las entidades espaciales del dominio de la aplicación pueden ser obtenidas a partir de la utilización este modelo.







Es decir, su estructura puede ser utilizada como un modelo o plantilla para crear un SIG censal. Al igual que con cualquier otro modelo, el usuario deberá adaptarlos a los requerimientos de su aplicación, lo cual implica que se pueden agregar nuevos atributos y operaciones, tanto espaciales como no-espaciales. El usuario puede, también, reemplazar o modificar los atributos y operaciones de las clases.

El modelo resultante es independiente de su implementación, esto permite que un componente de software, diseñado a partir de ella, pueda ser reutilizado.



## CONCLUSIONES

Esta investigación proporciona un punto de partida para resolver problemas censales que actualmente se filtran y comparan por métodos manuales. La investigación que llevamos a cabo para realizar el modelo de datos nos permitió profundizar sobre el paradigma orientado a objetos para el modelado de entidades geográficas.

Una vez planteado el modelo se debería poder utilizar para implementar un software que la aplique de la siguiente manera:

### *Precondiciones*

Los datos de los censos a consultar deberán estar cargados.

### *Datos de Entrada*

Consulta del usuario con los parámetros necesarios. Por ejemplo Crecimiento Poblacional de Capital Federal en los años 1990-2000 aplicando filtro de redondeo.

### *<Procesamiento de la consulta>*

### *Datos Salida*

#### Resultado del procesamiento de la consulta

El modelo captura las propiedades y comportamiento esencial de las entidades que intervienen en un censo poblacional tanto conceptuales como geográficas, de una manera más natural y simple que aquella empleada en la mayoría de herramientas SIG existentes, las cuales utilizan, por lo general, las representaciones vectorial o raster.

Más allá de su aplicación en el diseño conceptual de SIG y bases de datos espaciales, tiene utilidad en el desarrollo de componentes SIG reutilizables. En este sentido, el modelo que se plantea en este trabajo ha sido desarrollada con la idea de que pueda ser aplicada en la creación de marcos de aplicación dirigidos a facilitar la implementación de aplicaciones SIG orientadas a objetos.

Se plantea, también, la ampliación futura de la capacidad semántica del modelo a fin de que pueda ser útil para representar otros tipos de censos, tal y como lo describimos en el capítulo III, existen una variedad de censos (Poblacional, Agropecuario, etc.) y todos son necesarios en distintas áreas para la toma de decisiones.

## REFERENCIA BIBLIOGRÁFICA:

- [1] Aronoff
- [2] Star y estes
- [3] <http://www.GIS.sopde.es/cursosGIS/DHTML/que.html>
- [4] <http://www.nosoloGIS.com/index.asp>
- [5] <http://www.office.microsoft.com/latam/assistance/2002/articles/oUnderstandGraphicFormats.aspx>
- [6] [www.cesga.es/ca/defaultC.html?GIS/](http://www.cesga.es/ca/defaultC.html?GIS/)
- [7] Teoría de las poblaciones estables -Alfred Lotka (desarrollada en la primera mitad del siglo XX.)
- [8] <http://www.indec.mecon.gov.ar>
- [9] VINUESA, Julio (Ed.) (1994): Demografía. Análisis y Proyecciones. Síntesis.
- [10] Ortega Osona, José Antonio (2001): “Revisión de conceptos demográficos” en Contribuciones a la economía de La Economía de Mercado, virtudes e inconvenientes
- [11] <http://www.indec.mecon.ar/proyectos/censo2001/maestros/quees/masinfo.doc>
- [12] Marcela Fournari –Encargada de censistas - INDEC
- [13] Rumbaugh J., Jacobson J., Booch G., (1999): “The Unified Modeling Language Reference Manual”, Addison Wesley.
- [14]: Object-oriented software engineering: a use case driven approach I. Jacobson. Addison.- wesley 1992
- [15] M. Grand. Patterns in Java. Volume 1, John Wiley & Sons, New York,1998.
- [16] E Gamma; R. Helm; R. Johnson and J. Vlissides. Design patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, Reading, 1995.
- [17] B. Brugge and A.H. Dutoit. Object-Oriented Software Engineering. Prentice Hall, New Jersey, 2000.
- [18] M. Fowler. Software Patterns. En S. Zamir (ED.) Handbook of Object Technology. CRC Press, Florida, 1999, pags. 36.1-36.8
- [19] J. Soukup. Pattern Template Library. In S. Zamir (ed.), Handbook of Object Technology, CRC Press, Florida, 1999, pags. 37.1-37.18.
- [20] J. Bosque Sendra. Sistemas de Información Geográfica. Ediciones RIALP, Madrid,

1992.

[21] M.Wackowicz. Object-Oriented Design for Temporal GIS. Taylor&Francis, London, 1999.

[22]Patterns for GIS Applications Desing F.Balaguer, S.Gordillo,F.Das Neves.

[23] <http://www.inf.udec.cl/~mvaras/papers/2002/varas-urrutia-final.pdf>

[24] resarch article: “Can census Offices publish statistics form more than one samall area geography? An Analysis of differencing problem in statiscal disclosure”

[25] M. Fowler. Software Patterns. En S. Zamir (ED.) Handbook of Object Technology. CRC Press, Florida, 1999, pags. 36.1-36.8

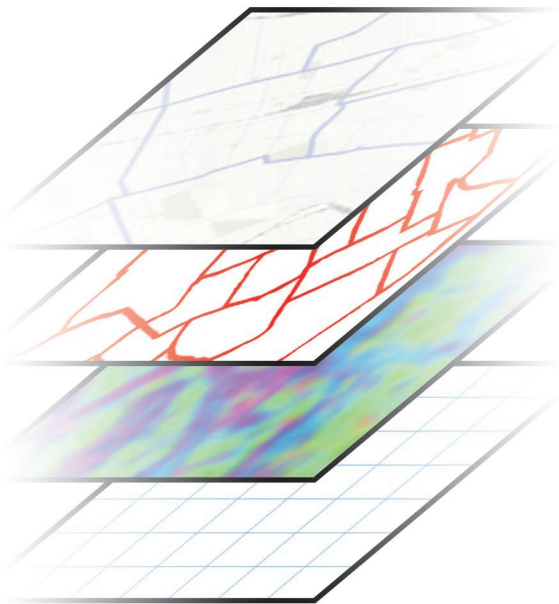
[26] Larman C. (1999), “UML y Patrones”, Prentice Hall.

[29] Poseydon CE - <http://www.gentleware.com>

[30]Towards a Patterns Language on Geographic Information Systems (GIS) Domain Vinko Vrsalovic Bolte, Yadrán Eterovic Solano.

Quiroga, Graciela Alejandra  
Ortiz Banda, Úrsula Italia

# DEFINICIÓN DE UN MODELO DE DISEÑO PARA CENSOS POBLACIONALES



Directora: Dra. Silvia Gordillo